T/AU00/00385

REC'D 16 MAY 2000

WIPO PCT

AU00/00385

Patent Office
Canberra

4

I, KAY WARD, TEAM LEADER EXAMINATION SUPPORT AND SALES hereby certify that annexed is a true copy of the Provisional specification in connection with Application No. PQ 0052 for a patent by COMMONWEALTH SCIENTIFIC AND INDUSTRIAL RESEARCH ORGANISATION, GOODMAN FIELDER LIMITED and GROUPE LIMAGRAIN PACIFIC PTY. LTD. filed on 29 April 1999.

COMMONWEALTH OF AUSTRALIA
AUSTRALIA
PATENT OFFICE

WITNESS my hand this
Eleventh day of May 2000

K Ward

KAY WARD
TEAM LEADER EXAMINATION
SUPPORT AND SALES

PRIORITY DOCUMENT

SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH RULE 17.1(a) OR (b)

Regulation 3.2

Commonwealth Scientific and Industrial Research Organisation AND Goodman Fielder Limited AND Group Limagrain Pacific Pty. Ltd.

AUSTRALIA

Patents Act 1990

PROVISIONAL SPECIFICATION

for the invention entitled:

"NOVEL GENES ENCODING WHEAT STARCH SYNTHASES AND USES THEREFOR"

The invention is described in the following statement:

NOVEL GENES ENCODING WHEAT STARCH SYNTHASES AND USES THEREFOR

FIELD OF THE INVENTION

The present invention relates generally to isolated nucleic acid molecules encoding wheat starch synthase enzymes and more particularly, to isolated nucleic acid molecules that encode wheat SSII and SSIII enzyme activities. The isolated nucleic acid molecules provide the means for modifying starch content and composition in plants, for example the ratio of amylose:amylopectin in the starch granule of the endosperm during the grain-filling phase of endosperm development. The isolated nucleic acid molecules of the present invention also provide the means for screening plant lines to determine the presence of natural and/or induced mutations in starch synthase genes which affect starch content and/or composition. The isolated nucleic acid molecules of the present invention further provide for the screening-assisted breeding of plants having desirable starch content and/or composition, in addition to providing for the direct genetic manipulation of plant starch content and/or composition.

GENERAL

Bibliographic details of the publications numerically referred to in this specification are collected at the end of the description.

This specification contains nucleotide and amino acid sequence information (SEQ ID Nos:) prepared using the programme PatentIn Version 2.0, presented herein at the end of the specification. Each nucleotide or amino acid sequence is identified in the sequence listing by the numeric indicator <210> followed by the sequence identifier (e.g. <210>1, <210>2, etc). The length, type of sequence (DNA, protein (PRT), etc) and source organism for each nucleotide or amino acid sequence are indicated by information provided in the numeric indicator fields <211>, <212> and <213>, respectively. Nucleotide and amino acid sequences (SEQ ID NOs:) referred to in the specification are defined by the information provided in numeric indicator field <400>

followed by the sequence identifier (eg. <400>1, <400>2, etc).

The designation of nucleotide residues referred to herein are those recommended by the IUPAC-IUB Biochemical Nomenclature Commission, wherein A represents Adenine, C represents Cytosine, G represents Guanine, T represents thymine, Y represents a pyrimidine residue, R represents a purine residue, M represents Adenine or Cytosine, K represents Guanine or Thymine, S represents Guanine or Cytosine, W represents Adenine or Thymine, H represents a nucleotide other than Guanine, B represents a nucleotide other than Adenine, V represents a nucleotide other than Thymine, D represents a nucleotide other than Cytosine and N represents any nucleotide residue.

The designations for naturally-occurring amino acid residues referred to herein are set forth in Table I. The designations for a non-limiting set of non-naturally-occurring amino acids is listed in Table 2.

As used herein the term "derived from" shall be taken to indicate that a specified integer may be obtained from a particular source albeit not necessarily directly from that source.

Throughout this specification, unless the context requires otherwise, the word "comprise", or variations such as "comprises" or "comprising", will be understood to imply the inclusion of a stated step or element or integer or group of steps or elements or integers but not the exclusion of any other step or element or integer or group of steps or elements or integers.

TABLE 1

Amino Acid	Three-letter Code	One-letter Code
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Cysteine	Cys	C
Glutamine	Gln	Q
Glutamic acid	Glu	E
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V
Aspartate/glutamate	Baa	B
Asparagine/glutamine		
Any amino acid as above	Xaa	X

TABLE 2

Non-conventional amino acid	Code	Non-conventional amino acid	Code
α-aminobutyric acid	Abu	L-N-methylalanine	Nmala
α-amino-α-methylbutyrate	Mgabu	L-N-methylarginine	Nmarg
aminocyclopropane-carboxylate	Cpro	L-N-methylasparagine	Nmasn
		L-N-methylaspartic acid	Nmasp
aminoisobutyric acid	Aib	L-N-methylcysteine	Nmcys
aminonorbornyl-carboxylate	Norb	L-N-methylglutamine	Nmgln
		L-N-methylglutamic acid	Nmglu
cyclohexylalanine	Chexa	L-N-methylhistidine	Nmhis
cyclopentylalanine	Cpen	L-N-methylisolleucine	Nmile
D-alanine	Dal	L-N-methylleucine	Nmleu
D-arginine	Darg	L-N-methyllysine	Nmlys
D-aspartic acid	Dasp	L-N-methylmethionine	Nmmet
D-cysteine	Dcys	L-N-methylnorleucine	Nmnle
D-glutamine	Dgln	L-N-methylnorvaline	Nmnva
D-glutamic acid	Dglu	L-N-methylornithine	Nmorn
D-histidine	Dhis	L-N-methylphenylalanine	Nmphe
D-isoleucine	Dile	L-N-methylproline	Nmpro
D-leucine	Dleu	L-N-methylserine	Nmser
D-lysine	Dlys	L-N-methylthreonine	Nmthr
D-methionine	Dmet	L-N-methyltryptophan	Nmtrp
D-ornithine	Dorn	L-N-methyltyrosine	Nmtyr
D-phenylalanine	Dphe	L-N-methylvaline	Nmval
D-proline	Dpro	L-N-methylethylglycine	Nmetg
D-serine	Dser	L-N-methyl-t-butylglycine	Nmtbug
D-threonine	Dthr	L-norleucine	Nle

D-tryptophan	Dtrp	L-norvaline	Nva
D-tyrosine	Dtyr	α-methyl-aminoisobutyrate	Maib
D-valine	Dval	α-methyl-γ-aminobutyrate	Mgabu
D-α-methylalanine	Dmala	α-methylcyclohexylalanine	Mchexa
D-α-methylarginine	Dmarg	α-methylcylcopentylalanine	Mcpen
D-α-methylasparagine	Dmasn	α-methyl-α-napthylalanine	Manap
D-α-methylaspartate	Dmasp	α-methylpenicillamine	Mpen
D-α-methylcysteine	Dmcys	N-(4-aminobutyl)glycine	Nglu
D-α-methylglutamine	Dmgln	N-(2-aminoethyl)glycine	Naeg
D-α-methylhistidine	Dmhis	N-(3-aminopropyl)glycine	Norn
D-α-methylisoleucine	Dmile	N-amino-α-methylbutyrate	Nmaabu
D-α-methylleucine	Dmleu	α-napthylalanine	Anap
D-α-methyllysine	Dmlys	N-benzylglycine	Nphe
D-α-methylmethionine	Dmmet	N-(2-carbamylethyl)glycine	Ngln
D-α-methylornithine	Dmorn	N-(carbamylmethyl)glycine	Nasn
D-α-methylphenylalanine	Dmphe	N-(2-carboxyethyl)glycine	Nglu
D-α-methylproline	Dmpro	N-(carboxymethyl)glycine	Nasp
D-α-methylserine	Dmser	N-cyclobutylglycine	Ncbut
D-α-methylthreonine	Dmthr	N-cycloheptylglycine	Nchep
D-α-methyltryptophan	Dmtrp	N-cyclohexylglycine	Nchex
D-α-methyltyrosine	Dmty	N-cyclodecylglycine	Ncdec
D-α-methylvaline	Dmval	N-cylcododecylglycine	Ncdod
D-N-methylalanine	Dnmala	N-cyclooctylglycine	Ncoct
D-N-methylarginine	Dnmarg	N-cyclopropylglycine	Ncpro
D-N-methylasparagine	Dnmasn	N-cycloundecylglycine	Ncund
D-N-methylaspartate	Dnmasp	N-(2,2-diphenylethyl) glycine	Nbhm
D-N-methylcysteine	Dnmcys	N-(3,3-diphenylpropyl) glycine	Nbhe

D-N-methylglutamine	Dnmgln	N-(3-guanidinopropyl) glycine	Narg
D-N-methylglutamate	Dnmglu	N-(1-hydroxyethyl)glycine	Nthr
D-N-methylhistidine	Dnmhis	N-(hydroxyethyl))glycine	Nser
D-N-methylisoleucine	Dnmile	N-(imidazolylethyl)) glycine	Nhis
D-N-methylleucine	Dnmleu	N-(3-indolylyethyl) glycine	Nhtrp
D-N-methyllysine	Dnmlys	N-methyl-γ-aminobutyrate	Nmgabu
N-methylcyclohexylalanine	Nmchexa	D-N-methylmethionine	Dnmmet
D-N-methylornithine	Dnmorn	N-methylcyclopentylalanine	Nmcpen
N-methylglycine	Nala	D-N-methylphenylalanine	Dnmphe
N-methylaminoisobutyrate	Nmaib	D-N-methylproline	Dnmpro
N-(1-methylpropyl)glycine	Nile	D-N-methylserine	Dnmser
N-(2-methylpropyl)glycine	Nleu	D-N-methylthreonine	Dnmthr
D-N-methyltryptophan	Dnmtrp	N-(1-methylethyl)glycine	Nval
D-N-methyltyrosine	Dnmtyr	N-methyla-napthylalanine	Nmanap
D-N-methylvaline	Dnmval	N-methylpenicillamine	Nmpen
γ-aminobutyric acid	Gabu	N-(*p*-hydroxyphenyl)glycine	Nhtyr
L-*t*-butylglycine	Tbug	N-(thiomethyl)glycine	Ncys
L-ethylglycine	Etg	penicillamine	Pen
L-homophenylalanine	Hphe	L-α-methylalanine	Mala
L-α-methylarginine	Marg	L-α-methylasparagine	Masn
L-α-methylaspartate	Masp	L-α-methyl-*t*-butylglycine	Mtbug
L-α-methylcysteine	Mcys	L-methylethylglycine	Metg
L-α-methylglutamine	Mgln	L-α-methylglutamate	Mglu
L-α-methylhistidine	Mhis	L-α-methylhomo phenylalanine	Mhphe
L-α-methylisoleucine	Mile	N-(2-methylthioethyl) glycine	Nmet

L-α-methylleucine	Mleu	L-α-methyllysine	Mlys
L-α-methylmethionine	Mmet	L-α-methylnorleucine	Mnle
L-α-methylnorvaline	Mnva	L-α-methylornithine	Morn
L-α-methylphenylalanine	Mphe	L-α-methylproline	Mpro
L-α-methylserine	Mser	L-α-methylthreonine	Mthr
L-α-methyltryptophan	Mtrp	L-α-methyltyrosine	Mtyr
L-α-methylvaline	Mval	L-N-methylhomo phenylalanine	Nmhphe
N-(N-(2,2-diphenylethyl) carbamylmethyl)glycine	Nnbhm	N-(N-(3,3-diphenylpropyl) carbamylmethyl)glycine	Nnbhe
1-carboxy-1-(2,2-diphenyl-ethylamino)cyclopropane	Nmbc		

Those skilled in the art will appreciate that the invention described herein is susceptible to variations and modifications other than those specifically described. It is to be understood that the invention includes all such variations and modifications. The invention also includes all of the steps, features, compositions and compounds referred to or indicated in this specification, individually or collectively, and any and all combinations or any two or more of said steps or features.

The present invention is not to be limited in scope by the specific embodiments described herein, which are intended for the purposes of exemplification only. Functionally-equivalent products, compositions and methods are clearly within the scope of the invention, as described herein.

BACKGROUND TO THE INVENTION

The biosynthesis of the starch granule is a complex process which involves the action of an array of isoforms of enzymes involved in the starch biosynthesis. Following the formation of glucose-1-phosphate, the enzyme activities required for the synthesis of granular starch include ADP glucose pyrophosphorylase (EC 2.7.7.27), starch synthases (EC 2.4.1.21), branching enzymes (EC 2.4.1.18) and debranching enzymes (EC 3.2.1.41 and EC 3.2.1.68) (Mouille *et al.*, 1996). Plants contain isozymes of each of these activities, and the definition of these isoforms and their roles has been conducted through investigation of the properties of the suite of soluble enzymes found in the stroma of the plastid, analysis of the proteins entrapped within the matrix of the starch granule, and mutational studies to identify genes and define linkages between individual genes and their specific roles.

Starch synthases extend regions of α-1,4 glucan through the transfer of the glucosyl moiety of ADPglucose to the non-reducing end of a pre-existing α-1,4 glucan. In addition to GBSS, 3 other classes of starch synthase have been identified in plants, SSI (wheat, Li *et al.*, 1999 and GenBank Accession No. U48227; rice, Baba *et al.*, 1993; potato, Genbank Accession No. STSTASYNT), SSII (pea, Dry *et al.* 1992; potato, Edwards *et al.*, 1995; maize, Harn *et al.* 1998 and GenBank Accession No. U66377) and SSIII (potato, Abel *et al.*, 1996; maize, Gao *et al.*, 1998). In the cereals, the most comprehensively studied species is maize, where in addition to GBSS, cDNAs encoding SSI, SSIIa, and SSIIb have been isolated, and both cDNA and genomic clones for *dull*1 have been characterised (Knight *et al.*, 1998; Harn *et al.*, 1998; Gao *et al.*, 1998). In maize, the product of the *du1* gene is known as maize SSII, however this gene is the homologue of potato SSIII.

The proteins within the matrix of the wheat starch granule have been extensively studied (Denyer *et al.*, 1995; Rahman *et al.*, 1995; Takaoka *et al.*, 1997; Yamamori and Endo, 1996) and 60, 75, 85, 100, 104 and 105 kDa protein bands can be visualised following SDS-PAGE. The predominant 60 kDa protein is exclusively

granule-bound and is analogous to the "waxy" granule bound starch synthase (GBSS) gene in maize (Rahman *et al.*, 1995). The combination of three null alleles for this enzyme from each of the wheat genomes (Nakamura *et al.*, 1995) results in the amylose-free "waxy" phenotype found in other species The 75 kDa starch synthase I (wSSI) is found in both the granule and the soluble fraction of wheat endosperm (Denyer *et al.*, 1995; Li *et al.*, 1999) and has been assigned to chromosomes 7A, 7B and 7D (Yamamori and Endo, 1996; Li *et al.*, 1999). The 85 kDa band contains a class II branching enzyme and an unidentified polypeptide (Rahman *et al.*, 1995). The 100, 104 and 105 kDa proteins of the wheat starch granule (designated Sgp-B1, Sgp-D1 and Sgp-A1 by Yamamori and Endo, 1996) have been shown to be encoded by a homeologous set of genes on the short arm of chromosome 7B, 7A and 7D respectively (Yamamori and Endo, 1996; Takaoka *et al.*, 1997). Denyer *et al.* (1995) concluded on the basis of enzyme activity assays that these proteins were also starch synthases. These genes are referred to hereinafter as the "wheat SSII genes".

While GBSS has been established to be essential for amylose synthesis, the remaining starch synthases are thought to be primarily responsible for the elongation of amylopectin chains, although this does not preclude them from also having non-essential roles in amylose biosynthesis. Differences in kinetic properties between isoforms, and the analysis of mutants lacking various isoforms, suggests that each isoenzyme contributes to the extension of specific subsets of the available non-reducing ends. Accordingly, the production of plants that produce improved starches that are modified for particular purposes, for example starches having high or low amylose:amylopectin ratios, requires the availability of genes encoding the various starch synthase isoforms. Moreover, because of species-specific codon usages and variations in the kinetic parameters of these isoforms in different species, the production of modified starches may require the use of genes derived from particular species.

In work leading up to the present invention, the inventors sought to modify wheat

starch composition and content, by providing isolated nucleotide sequences encoding the wheat SSII (i.e. wSSII) and wheat SSIII (i.e. wSSIII) isoenzymes, and by introducing these nucleotide sequences into plants using recombinant DNA technology.

SUMMARY OF THE INVENTION

The present invention provides isolated nucleic acid molecules encoding the 100, 104 and 105 kDa SSII (Sgp-1) polypeptides of the wheat starch granule matrix, as determined using the SDS/PAGE system of Rahman *et al.* (1995), which polypeptides are equivalent to the 100, 108 and 115 kDa polypeptides described by Yamamori and Endo (1996). The present invention further provides isolated nucleic acid molecules encoding the soluble *dull*1-type wheat starch synthase III polypeptide. Analysis of the polypeptides encoded by these nucleic acid molecules reveals several consensus amino acid sequence motifs (i.e., sequences having at least 25% sequence identity to any one or more of the amino acid sequences selected from the group consisting of (a)KVGGLGDVVTS;(b)GHTVEVILPKY;(c) HDWSSAPVAWLYKEHY; (d) GILNGIDPDIWDPYTD; (e) DVPIVGIITRLTAQKG; (f)NGQVVLLGSA; (g)AGSDFIIVPSIFEPCGLTQLVAMRYGS; and (h)TGGLVDTV) that are highly conserved in wheat starch synthase isoenzymes, in addition to isoenzyme-specific sequences, which sequences possess utility in isolating related starch synthase-encoding sequences and in assaying plants for their expression of one or more starch synthase isoenzymes.

Accordingly, one aspect of the present invention provides an isolated nucleic acid molecule which comprises a sequence of nucleotides which encodes, or is complementary to a nucleic acid molecule which encodes a wheat starch synthase polypeptide, protein or enzyme molecule or a functional subunit thereof selected from the following:

(i) a wheat starch synthase II (wSSII) polypeptide, protein or enzyme or functional subunit thereof which comprises an amino acid sequence which is at

least about 85% identical overall to an amino acid sequence set forth in any one of SEQ ID NOS: <400>2, <400>4, or <400>6;

(ii) a wheat starch synthase III (wSSIII) polypeptide, protein or enzyme or functional subunit thereof which comprises an amino acid sequence which is at least about 85% identical overall to an amino acid sequence set forth in any one of SEQ ID NOS: <400>8 or <400>10;and

(iii) a wheat starch synthase polypeptide, protein or enzyme or functional subunit thereof which comprises a conserved amino acid sequence having at least 25% identity to an amino acid sequence selected from the group consisting of:

(a) KVGGLGDVVTS;

(b) GHTVEVILPKY;

(c) HDWSSAPVAWLYKEHY;

(d) GILNGIDPDIWDPYTD;

(e) DVPIVGIITRLTAQKG;

(f) NGQVVLLGSA;

(g)AGSDFIIVPSIFEPCGLTQLVAMRYGS; and

(h)TGGLVDTV

and wherein said wheat starch synthase polypeptide further comprises an amino acid sequence having at least about 85% identity overall to an amino acid sequence set forth in any one of SEQ ID NOS: <400>2, <400>4, <400>6, <400>8 or <400>10.

In a preferred embodiment, the isolated nucleic acid molecule encodes a starch synthase polypeptide, protein or enzyme having at least about 90% amino acid sequence identity to any one of SEQ ID NOS:<400>2, <400>4, <400>6, <400>8 or <400>10, more preferably having at least about 95% or about 97% or about 99% identity to any one of said amino acid sequences.

In an alternative embodiment, the present invention provides an isolated nucleic acid

molecule which encodes a wheat starch synthase polypeptide, protein or enzyme molecule or a functional subunit thereof, wherein said nucleic acid molecule comprises a nucleotide sequence having at least about 85% nucleotide sequence identity to any one of SEQ ID NOS: <400>1, <400>3, <400>5, <400>7 or <400>9 or SEQ ID NOS: <400>11 to <400>16, or a complementary nucleotide sequence thereto.

In a preferred embodiment, the isolated nucleic acid molecule comprises the nucleotide sequence set forth in any one of SEQ ID NOS: <400>1, <400>3, <400>5, <400>7 or <400>9 or SEQ ID NOS: <400>11 to <400>16, or is at least about 90% identical, more preferably at least about 95% or 97% or 99% identical to all or a protein-encoding part thereof.

In an alternative embodiment, the present invention provides an isolated nucleic acid molecule which encodes a wheat starch synthase polypeptide, protein or enzyme molecule or a functional subunit thereof, wherein said nucleic acid molecule comprises a nucleotide sequence that is capable of hybridising under at least moderate stringency hybridisation conditions to at least about 30 contiguous nucleotides derived from any one of SEQ ID NOS: <400>1, <400>3, <400>5, <400>7 or <400>9 or SEQ ID NOS: <400>11 to <400>16, or a complementary nucleotide sequence thereto.

A second aspect of the present invention provides a method of isolating a nucleic acid molecule that encodes a starch synthase polypeptide, protein or enzyme having at least about 85% amino acid sequence identity to any one SEQ ID NOS:<400>2, <400>4, <400>6, <400>8 or <400>10 and/or which comprises anamino acid sequence having at least 25% identity to an amino acid sequence selected from the group consisting of:

(a) KVGGLGDVVTS;

(b) GHTVEVILPKY;

(c) HDWSSAPVAWLYKEHY;

(d) GILNGIDPDIWDPYTD;

(e) DVPIVGIITRLTAQKG;

(f) NGQVVLLGSA;

(g)AGSDFIIVPSIFEPCGLTQLVAMRYGS; and

(h)TGGLVDTV ,

said method comprising:

(i) hybridising a probe or primer comprising at least about 15 contiguous nucleotides in length derived from any one of SEQ ID NOS: <400>1, <400>3, <400>5, <400>7 or <400>9 or SEQ ID NOS: <400>11 to <400>16, or a complementary nucleotide sequence thereto to single-stranded or double-stranded mRNA, cDNA or genomic DNA; and

(ii) detecting the hybridised mRNA, cDNA or genomic DNA using a detecting means.

Preferably, the detecting means is a reporter molecule covalently attached to the probe or primer molecule or alternatively, a polymerase chain reaction format. Accordingly, the present invention clearly extends to the use of the nucleic acid molecules provided herein to isolate related starch synthase-encoding sequences using standard hybridisation and/or polymerase chain reaction techniques.

A third aspect of the invention provides an isolated probe or primer comprising at least about 15 contiguous nucleotides in length derived from any one of SEQ ID NOS: <400>1, <400>3, <400>5, <400>7 or <400>9 or SEQ ID NOS: <400>11 to <400>16, or a complementary nucleotide sequence thereto.

Preferably, the probe or primer comprises a nucleotide sequence set forth in any one of SEQ ID NOS:<400>25 to <400>34.

A fourth aspect of the present invention is directed to an isolated or recombinant starch synthase polypeptide, protein or enzyme, preferably substantially free of conspecific or non-specific proteins, which comprises an amino acid sequence selected from the

following:

(i) a wheat starch synthase II (wSSII) polypeptide, protein or enzyme or functional subunit thereof which comprises an amino acid sequence which is at least about 85% identical overall to an amino acid sequence set forth in any one of SEQ ID NOS: <400>2, <400>4, or <400>6;

(ii) a wheat starch synthase III (wSSIII) polypeptide, protein or enzyme or functional subunit thereof which comprises an amino acid sequence which is at least about 85% identical overall to an amino acid sequence set forth in any one of SEQ ID NOS: <400>8 or <400>10;and

(iii) a wheat starch synthase polypeptide, protein or enzyme or functional subunit thereof which comprises a conserved amino acid sequence having at least 25% identity to an amino acid sequence selected from the group consisting of:

(a) KVGGLGDVVTS;
(b) GHTVEVILPKY;
(c) HDWSSAPVAWLYKEHY;
(d) GILNGIDPDIWDPYTD;
(e) DVPIVGIITRLTAQKG;
(f) NGQVVLLGSA;
(g)AGSDFIIVPSIFEPCGLTQLVAMRYGS; and
(h)TGGLVDTV

and wherein said wheat starch synthase polypeptide further comprises an amino acid sequence having at least about 85% identity overall to an amino acid sequence set forth in any one of SEQ ID NOS: <400>2, <400>4, <400>6, <400>8 or <400>10.

A further aspect of the invention provides a method of assaying for the presence or absence of a starch synthase isoenzyme or the copy number of a gene encoding same in a plant, comprising contacting a biological sample derived from said plant with an

isolated nucleic acid molecule derived from any one of SEQ ID NOS: <400>1, <400>3, <400>5, <400>7 or <400>9 or SEQ ID NOS: <400>11 to <400>16, or any one of SEQ ID NOS:<400>25 to <400>34, or a complementary nucleotide sequence thereto for a time and under conditions sufficient for hybridisation to occur and then detecting said hybridisation using a detection means.

The detection means according to this aspect of the invention is any nucleic acid based hybridisation or amplification reaction.

A further aspect of the present invention utilises the above-mentioned assay method in the breeding and/or selection of plants which express or do not express particular starch sythase isoenzymes or alternatively, which express a particular starch synthase isoenzyme at a particular level in one or more plant tissues. This aspect clearly extends to the selection of transformed plant material which contains one or more of the isolated nucleic acid molecules of the present invention.

A further aspect of the present invention provides a method of modifying the starch content and/or starch composition of one or more tissues or organs of a plant, comprising expressing therein a sense molecule, antisense molecule, ribozyme molecule, co-suppression molecule, or gene-targeting molecule having at least about 85% nucleotide sequence identity to any one of any one of SEQ ID NOS: <400>1, <400>3, <400>5, <400>7 or <400>9 or SEQ ID NOS: <400>11 to <400>16, or a complementary nucleotide sequence thereto for a time and under conditions sufficient for the enzyme activity of one or more starch synthase isoenzymes to be modified. This aspect of the invention clearly extends to the introduction of the sense molecule, antisense molecule, ribozyme molecule, co-suppression molecule, or gene-targeting molecule to isolated plant cells, tissues or organs or organelles by cell fusion or transgenic means and the regeneration of intact plants therefrom.

A further aspect of the present invention provides an isolated promoter that is operable

in the endosperm of a monocotyledonous plant cell, tissue or organ, and preferably in the endosperm of a monocotyledonous plant cell, tissue or organ. For example, the HMG promoter from wheat, or the maize zein gene promoter are particularly preferred, as is the promoter derived from a starch synthase gene of the present invention, such as a promoter that is linked *in vivo* to any one of SEQ ID NOS: <400>1, <400>3, <400>5, <400>7 or <400>9 or any one of SEQ ID NOS: <400>11 to <400>16, or a complementary nucleotide sequence thereto.

A still further aspect of the present invention contemplates a transgenic plant comprising an introduced sense molecule, antisense molecule, ribozyme molecule, co-suppression molecule, or gene-targeting molecule having at least about 85% nucleotide sequence identity to any one of any one of SEQ ID NOS: <400>1, <400>3, <400>5, <400>7 or <400>9 or SEQ ID NOS: <400>11 to <400>16, or a complementary nucleotide sequence thereto or a gentic construct comprising same, and to plant propagules, cells, tissues, organs or plant parts derived from said transgenic plant that also carry the introduced molecule(s).

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is a copy of a photographic representation showing the distribution of wheat endosperm starch synthases between the starch granule and soluble fractions. Lane 1, SDS-PAGE of wheat endosperm starch granule proteins revealed by silver staining; lanes 2-7, immunoblot of wheat endosperm soluble phase and starch granule proteins separated by SDS-PAGE from various developmental stages and probed with an anti-(wheat wSSII peptide) monoclonal antibody. Lanes 2-4 contain proteins from the soluble fraction of wheat endosperm at 15 days post anthesis (Lane 2); 20 days post anthesis (Lane 3); and at 25 days post anthesis (Lane 4). Lanes 5-7 contain proteins from the starch granule of wheat endosperm at 15 days post anthesis (Lane 5); 20 days post anthesis (Lane 6); and at 25 days post anthesis (Lane 7).

Figure 2 is a copy of a schematic representation comparing the nucleotide sequences

of cDNA clones designated wSSIIA, wSSIIB and wSSIID, encoding the starch synthase II polypeptides from wheat, using the PILEUP programme of Devereaux *et al.* (1984).

Figure 3 is a copy of a schematic representation comparing the deduced amino acid sequences of starch synthase II from wheat (wSSIIA, wSSIIB and wSSIID), maize (maizeSSIIa and maizeSSIIb; Harn *et al.*, 1998), pea (peaSSII; Dry et al ., 1992) and potato (potatoSSII; van der Leij *et al.*, 1991). Identical amino acid residues among each of these sequences are indicated below the sequences with "*". The alignments of maize SSIIa with maize SSIIb, and pea SSII and potato SSII are essentially as described in Harn *et al.* (1998) and Edwards *et al.* (1995). All sequences are aligned to position the transit peptide cleavage site below the arrow (⇓) between residues 59 and 60 of the wSSIIA sequence. The wSSIIp1 sequence, the sequence of SGP-B1 (peptide3), and of eight conserved regions are annotated and underlined.

Figure 4 is a copy of a photographic representation of a northern blot showing the expression of wheat wSSII mRNA in wheat plants. Total RNAs were isolated from leaves pre-anthesis florets and endosperm of the wheat cultivar "Gabo", grown under a photoperiod comprising 16 hours daylength, and at 18 °C during theday, and at 13 °C during the night cycle, and probed with the wSSIIp2 DNA fragment. The source of each RNA is indicated at the top of the Figure as follows: Lane 1, leaf; Lane 2, pre-anthesis florets; Lanes 3-11, endosperm at: 4 days post-anthesis (Lane 3); 6 days post-anthesis (Lane 4); 8 days post-anthesis (Lane 5); 10 days post-anthesis (Lane 6);12 days post-anthesis (Lane 7); 15 days post-anthesis (Lane 8); 18 days post-anthesis (Lane 9); 21 days post-anthesis (Lane 10); and 25 days post-anthesis (Lane 11).

Figure 5 is a copy of a photographic representation showing the localization of wheat starch synthase II genes on the wheat genome by PCR, using the primers ssIIc, ssIId and ssIIe in the amplification reaction. The nullisomic-tetrasomic genomic DNA of

wheat cv. Chinese Spring was used as template DNA. Lane D, *Triticum tauschii*; Lane AB, Accession line N7DT7B having no 7D chromosome and four copies of the 7B chromosome; Lane AD, Accession line N7BT7A having no 7B chromosome and four copies of the 7A chromosome; Lane BD, Accession line N7AT7B having no 7A chromosome and four copies of the 7B chromosome; Lane ABD, wheat cv. Chinese Spring. PCR products derived from each cDNA clone are labelled. The results indicate that the cDNA clones, wSSIIB, wSSIIA and wSSIID are derived from the B-, A- and D-genomes of wheat, respectively.

Figure 6 is a copy of a photographic representation showing the purification of a wheat SSII genomic clone from the *T. tauschii* var. Strangulata (Accession No. CPI 110799) genomic library. A genomic clone, designated wSSII-8, was identified by hybridisation with the wSSIIp2 probe and purified through successive rounds of selection and hybridisation.

Figure 7 is a copy of a photographic representation showing a Southern blot of *Bam*HI-digested genomic clone DNAs identified in a primary screening using the wSSIIp2 probe, following hybridisation with wSSIIp4 probe DNA which is derived from the 5′-end of the wSSIIA cDNA clone. Lane 8 contains DNA derived from genomic clone wSSII-8 (see Figure 6). Hybridisation of clone wSSII-8 to thw wSSIIp4 probe suggests that this genomic clone contains the promoter region of the wSSII gene.

Figure 8 is a schematic representation comparing the deduced amino acid Sequences of the maize, potato and wheat SSIII polypeptides.

Figure 9 is a copy of a photographic representation showing the purification of a wheat SSIII genomic clone from a *T. tauschii* genomic library. A plaque was identified by hybridisation with a PCR-derived from the wSSIII.B3 gene (a) and purified through successive rounds of selection and hybridisation. The hybridisation of plaques from a third round of plaque purification is shown in (b).

Figure 10 is a copy of a photographic representation showing the expression of wheat wSSIII mRNA in wheat. Total RNAs were isolated from the endosperm of the wheat cultivars Wyuna (Panel a) and Gabo (Panel b) leaves pre-anthesis florets and endosperm of the wheat cultivar "Gabo", grown under a photoperiod comprising 16 hours daylength, and at 18 ^{o}C during the day cycle, and at 13 ^{o}C during the night cycle, and probed with the wSSIIIp1 DNA fragment derived from wSSIII.B3 cDNA. The source of each RNA is indicated at the top of the Figure as follows: Lane 1, endosperm at: 4 days post-anthesis; Lane 2, endosperm at 6 days post-anthesis; Lane 4, endosperm at 8 days post-anthesis; Lane 4, endosperm at 10 days post-anthesis; Lane 5, endosperm at 12 days post-anthesis; Lane 6, endosperm at 15 days post-anthesis; Lane 7, endosperm at 18 days post-anthesis; Lane 8, endosperm at 21 days post-anthesis; Lane 9, endosperm at 25 days post-anthesis; and Lane 10, endosperm at 31 days post-anthesis (Panel a only). In panel (c), L refers to leaf RNA, and P refers to RNA from pre-anthesis florets derived from the cultivar Gabo.

Figure 11 is a schematic representation showing the relationships between the primary amino acid sequences of starch synthases (SS) and glycogen synthase of *E. coli* (GS). The dendrogram was generated by the program PILEUP (Devereaux *et al.*, 1984). The amino acid sequences used for the analysis are those of the wheat SSIIA, wheat SSIIB, wheat SSIID, and wheat SSIII polypeptides of the present invention compared to the deduced amino acid sequences of wheat GBSS (Clark *et al.*, 1991), wheat SSI (Li *et al.*, 1999), rice GBSS (Okagaki, 1992), rice SSI (Baba *et al.*, 1993), maize GBSS (Kloesgen *et al.*, 1986), maize SSI (Knight *et al.*, 1998), maize SSIIa and maize SSIIb (Harn *et al.*, 1998), maize SSIII (Gao *et al.*, 1998), pea GBSS (Dry *et al.*, 1992), pea SSII (Dry *et al.*, 1992), potato GBSS (van der Leij *et al.*, 1991), potato SSI (Genbank accession number: STSTASYNT), potato SSII (Edwards *et al.*, 1995), potato SSIII (Abel *et al.*, 1996), and *E. coli* glycogen synthase (GS) (Kumar *et al.*, 1986). Five groups of enzymes included in the alignment are granule-bound starch synthase (GBSS), starch synthase-I (SSI), starcg synthase-II (SSII), starch synthase-III (SSIII) and glycogen synthase (GS).

Figure 12 is a schematic representation showing the position of conserved regions within cereal starch synthase genes. Comparisons of cereal starch synthases were made based on their deduced amino acid sequences and 8 conserved regions identified. Conserved regions are shown in bold and transit peptides (where defined) in grey. The sequences included in the alignment are the wheat SSII-A1 and wheat SSIII polypeptides of the present invention; wheat GBSS (Ainsworth *et al.*, 1993); wheat SSI (Li *et al.*, 1999); maize SSIIa (Harn *et al.*, 1998); and maize dull-1(Gao *et al.*, 1998).

Figure 13 is a schematic representation showing the position of conserved amino acid seqeunces within four wheat starch synthase proteins. The eight highly-conserved regions between the wheat starch synthase polypeptides are underlined and annotated at the top of each group of amino acid sequences. The sequences included in the alignment are the wheat SSII-A1 and wheat SSIII polypeptides of the present invention; wheat GBSS (wGBSS; Yan *et al.*, 1999); wheat SSI (wSS1; Li *et al.*, 1999); wheat SSII (wSS2; SEQ ID NO:<400>4); and wheat SSIII (wSS3; SEQ ID NO:<400>8).

Figure 14 is a copy of a schematic representation of a gene map showing the alignment of fragments 1 to 6 of the genomic SSIII gene (lower line) with the corresponding SSIII cDNA clone (uper line). Raised regions in the genomic clone fragments (lower line) represent protein-encoding regions of the gene.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

One aspect of the present invention provides an isolated nucleic acid molecule which comprises a sequence of nucleotides which encodes, or is complementary to a nucleic acid molecule which encodes a wheat starch synthase polypeptide, protein or enzyme molecule or a functional subunit thereof selected from the following:

(i) a wheat starch synthase II (wSSII) polypeptide, protein or enzyme or functional subunit thereof which comprises an amino acid sequence set forth

in any one of SEQ ID NOS: <400>2, <400>4, or <400>6; and

(ii) a wheat starch synthase III (wSSIII) polypeptide, protein or enzyme or functional subunit thereof which comprises an amino acid sequence set forth in any one of SEQ ID NOS: <400>8 or <400>10.

Alternatively or in addition, the isolated nucleic acid molecule of the present invention encodes a wheat starch synthase II (wSSII) polypeptide, protein or enzyme or functional subunit thereof and comprises a nucleotide sequence set forth in any one of SEQ ID NOS: <400>1, <400>3, or <400>5.

Alternatively or in addition, the isolated nucleic acid molecule of the present invention encodes a wheat starch synthase III (wSSIII) polypeptide, protein or enzyme or functional subunit thereof and comprises a nucleotide sequence set forth in any one of SEQ ID NOS: <400>7 or <400>9.

As used herein, the term "starch synthase" shall be taken to refer to any enzymatically-active peptide, polypeptide, oligopeptide, polypeptide, protein or enzyme molecule that is at least capable of transferring a glucosyl moiety from ADP-glucose to an α-1,4-glucan molecule, or a peptide, polypeptide, oligopeptide or polypeptide fragment of such an enzymatically-active molecule.

The term "wheat starch synthase" refers to a starch synthase derived from hexaploid wheat or barley or a progenitor species, or a relative thereto such as the diploid *Triticum tauschii* or other diploid, tetraploid, aneuploid, polyploid, nullisomic, or a wheat/barley addition line, amongst others, the only requirement that the genomic DNA is at least about 80% identical to the genome of a wheat plant as determined by standard DNA melting curve analyses.

The term "starch synthase II" or "wSSII" or similar term shall be taken to refer to a starch synthase as hereinbefore defined that is detectable in the starch granule of a

plant seed endosperm and possesses one or more properties selected from the group consisting of:

(i) it is immunologically cross-reactive with the wheat starch granule proteins designated Sgp-B1 and/or Sgp-D1 and/or Sgp-A1, having estimated molecular weights of about 85 kDa to about 115 kDa;

(ii) it is encoded by one of a homeologous set of genes localised on wheat chromosomes 7B or 7A or 7D;

(iii) it is encoded by a nucleotide sequence that comprises at least about 15 nucleotides in length derived from any one or more of SEQ ID NOS: <400>1, <400>3, or <400>5 or a complementary nucleotide sequence thereto;

(iv) it is encoded by a nucleotide sequence that is at least about 85% identical to one or more of the nucleotide sequences set forth in SEQ ID NOS: <400>1, <400>3, or <400>5 or a complementary nucleotide sequence thereto;

(v) it comprises an amino acid sequence having at least about 85% identity to one or more of SEQ ID NOS:<400>2 or <400>4 or <400>6;

(vi) it comprises at least about 5 contiguous amino acids, preferably at least about 10 contiguous amino acids, more preferably at least about 15 contiguous amino acids, even more preferably at least about 20 contiguous amino acids and still even more preferably at least about 25-50 contiguous amino acids of the amino acid sequences set forth in SEQ ID NOS:<400>2 or <400>4 or <400>6;and

(vii) it which comprises a conserved amino acid sequence having at least 25% identity to an amino acid sequence selected from the group consisting of:

(a) KVGGLGDVVTS;

(b) GHTVEVILPKY;

(c) HDWSSAPVAWLYKEHY;

(d) GILNGIDPDIWDPYTD;

(e) DVPIVGIITRLTAQKG;

(f) NGQVVLLGSA;

(g)AGSDFIIVPSIFEPCGLTQLVAMRYGS; and

(h)TGGLVDTV

in addition to any one or more of (i) to (vi).

The term "starch synthase III" or "wSSIII" or similar term shall be taken to refer to a starch synthase as hereinbefore defined that possesses one or more properties selected from the group consisting of:

(i) it is encoded by a nucleotide sequence that comprises at least about 15 nucleotides in length derived from any one or more of SEQ ID NOS:<400>7 or <400>9 or any one or more of SEQ ID NOS: <400>11 to <400>16 or a complementary nucleotide sequence thereto;

(ii) it is encoded by a nucleotide sequence that is at least about 85% identical to one or more of the nucleotide sequences set forth in SEQ ID NOS: <400>7 or <400>9 or any one or more of SEQ ID NOS: <400>11 to <400>16 or a complementary nucleotide sequence thereto; and

(iii) it comprises an amino acid sequence having at least about 85% identity to one or more of SEQ ID NOS:<400>8 or <400>10;

(iv) it comprises at least about 5 contiguous amino acids, preferably at least about 10 contiguous amino acids, more preferably at least about 15 contiguous amino acids, even more preferably at least about 20 contiguous amino acids and still even more preferably at least about 25-50 contiguous amino acids of the amino acid sequences set forth in SEQ ID NOS:<400>8 or <400>10;and

(v) which comprises a conserved amino acid sequence having at least 25% identity to an amino acid sequence selected from the group consisting of:

(a) KVGGLGDVVTS;

(b) GHTVEVILPKY;

(c) HDWSSAPVAWLYKEHY;

(d) GILNGIDPDIWDPYTD;

(e) DVPIVGIITRLTAQKG;

(f) NGQVVLLGSA;

(g)AGSDFIIVPSIFEPCGLTQLVAMRYGS; and

(h)TGGLVDTV

in addition to any one or more of (i) to (iv).

In a more preferred embodiment, the WSSII or WSSIII polypeptide encoded by the nucleic acid molecule of the present invention will comprise a substantial contiguous region of any one of SEQ ID NOS: <400>2, <400>4, <400>6, <400>8 or <400>10 or <400>17 sufficient to possess the biological activity of a starch synthase polypeptide.

For the purposes of nomenclature, the nucleotide sequence set forth in SEQ ID NO: <400>1 relates to the cDNA molecule encoding the WSSII (i.e. Sgp-B1) polypeptide of wheat. The amino acid sequence of the corresponding polypeptide is set forth herein as SEQ ID NO:<400>2. The nucleotide sequence set forth in SEQ ID NO: <400>3 relates to the cDNA molecule encoding the WSSII (i.e. Sgp-A1) polypeptide of wheat. The amino acid sequence of the corresponding polypeptide is set forth herein as SEQ ID NO:<400>4. The nucleotide sequence set forth in SEQ ID NO: <400>5 relates to the cDNA molecule encoding the WSSII (i.e. Sgp-D1) polypeptide of wheat. The amino acid sequence of the corresponding polypeptide is set forth herein as SEQ ID NO:<400>6. The nucleotide sequences set forth in SEQ ID NOs: <400>7 and <400>9 relate, respectively, to full-length and partial cDNA molecules encoding the WSSIII polypeptide of wheat. The amino acid sequences of the corresponding polypeptides are set forth herein as SEQ ID NOs:<400>8 and <400>10, respectively.The nucleotide sequences set forth in SEQ ID NOs: <400>11 to <400>16 relates to fragments of the genomic gene encoding the WSSIII polypeptide of wheat, significant protein-encoding regions of which are described by reference to Table 3 and Figure 14.

Preferably, the isolated nucleic acid molecule of the present invention comprises a sequence of nucleotides which encodes, or is complementary to a nucleic acid molecule which encodes a wheat starch synthase III polypeptide, protein or enzyme molecule or a functional subunit thereof which comprises an amino acid sequence

which is at least about 85% identical overall to an amino acid sequence set forth in any one of SEQ ID NOS: <400>8 or <400>10 and more preferably, which additionally comprises which comprises one or more conserved amino acid sequences selected from the group consisting of:

(a) KVGGLGDVVTS;
(b) GHTVEVILPKY;
(c) HDWSSAPVAWLYKEHY;
(d) GILNGIDPDIWDPYTD;
(e) DVPIVGIITRLTAQKG;
(f) NGQVVLLGSA;
(g)AGSDFIIVPSIFEPCGLTQLVAMRYGS; and
(h)TGGLVDTV .

The present invention clearly extends to homologues, analogues and derivatives of the wheat starch synthase II and III genes exemplified by the nucleotide sequences set forth herein as SEQ ID NOs:<400>1, <400>3, <400>5, <400>7, <400>9 and <400>11 to <400>16.

Preferred starch synthase genes may be derived from a naturally-occurring starch synthase gene by standard recombinant techniques. Generally, a starch synthase gene may be subjected to mutagenesis to produce single or multiple nucleotide substitutions, deletions and/or additions. Nucleotide insertional derivatives of the starch synthase gene of the present invention include 5′ and 3′ terminal fusions as well as intra-sequence insertions of single or multiple nucleotides. Insertional nucleotide sequence variants are those in which one or more nucleotides are introduced into a predetermined site in the nucleotide sequence although random insertion is also possible with suitable screening of the resulting product. Deletional variants are characterised by the removal of one or more nucleotides from the sequence. Substitutional nucleotide variants are those in which at least one nucleotide

in the sequence has been removed and a different nucleotide inserted in its place. Such a substitution may be "silent" in that the substitution does not change the amino acid defined by the codon. Alternatively, substituents are designed to alter one amino acid for another similar acting amino acid, or amino acid of like charge, polarity, or hydrophobicity.

For the present purpose, "homologues" of a nucleotide sequence shall be taken to refer to an isolated nucleic acid molecule which is substantially the same as the nucleic acid molecule of the present invention or its complementary nucleotide sequence, notwithstanding the occurrence within said sequence, of one or more nucleotide substitutions, insertions, deletions, or rearrangements.

"Analogues" of a nucleotide sequence set forth herein shall be taken to refer to an isolated nucleic acid molecule which is substantially the same as a nucleic acid molecule of the present invention or its complementary nucleotide sequence, notwithstanding the occurrence of any non-nucleotide constituents not normally present in said isolated nucleic acid molecule, for example carbohydrates, radiochemicals including radionucleotides, reporter molecules such as, but not limited to DIG, alkaline phosphatase or horseradish peroxidase, amongst others.

"Derivatives" of a nucleotide sequence set forth herein shall be taken to refer to any isolated nucleic acid molecule which contains significant sequence similarity to said sequence or a part thereof. Generally, the nucleotide sequence of the present invention may be subjected to mutagenesis to produce single or multiple nucleotide substitutions, deletions and/or insertions. Nucleotide insertional derivatives of the nucleotide sequence of the present invention include 5′ and 3′ terminal fusions as well as intra-sequence insertions of single or multiple nucleotides or nucleotide analogues. Insertional nucleotide sequence variants are those in which one or more nucleotides or nucleotide analogues are introduced into a predetermined site in the nucleotide sequence of said sequence, although random insertion is also possible with suitable

screening of the resulting product being performed. Deletional variants are characterised by the removal of one or more nucleotides from the nucleotide sequence. Substitutional nucleotide variants are those in which at least one nucleotide in the sequence has been removed and a different nucleotide or nucleotide analogue inserted in its place.

The present invention extends to the isolated nucleic acid molecule when integrated into the genome of a cell as an addition to the endogenous cellular complement of starch synthase genes, irrespective of whether or not the introduced nucleotide sequence is translatable or non-translatable to produce a polypeptide. The present invention clearly contemplates the introduction of additional copies of starch synthase genes into plants, particularly wheat plants, in the antisense orientation to reduce the expression of particular wheat starch synthase genes. As will be known to those skilled in the art, such antisense genes are non-translatable, notwithstanding that they can be expressed to produce antisense mRNA molecules.

The said integrated nucleic acid molecule may, or may not, contain promoter sequences to regulate expression of the subject genetic sequence.

Accordingly, the present invention clearly encompasses preferred homologues, analogues and derivatives that comprise a sequence of nucleotides which encodes, or is complementary to a nucleic acid molecule which encodes a wheat starch synthase polypeptide, protein or enzyme molecule or a functional subunit thereof selected from the following:

(i) a wheat starch synthase II (wSSII) polypeptide, protein or enzyme or functional subunit thereof which comprises an amino acid sequence which is at least about 85% identical overall to an amino acid sequence set forth in any one of SEQ ID NOS: <400>2, <400>4, or <400>6;

(ii) a wheat starch synthase III (wSSIII) polypeptide, protein or enzyme or functional subunit thereof which comprises an amino acid sequence which is at

least about 85% identical overall to an amino acid sequence set forth in any one of SEQ ID NOS: <400>8 or <400>10;and

(iii) a wheat starch synthase polypeptide, protein or enzyme or functional subunit thereof which comprises a conserved amino acid sequence having at least 25% identity to an amino acid sequence selected from the group consisting of:

(a) KVGGLGDVVTS;
(b) GHTVEVILPKY;
(c) HDWSSAPVAWLYKEHY;
(d) GILNGIDPDIWDPYTD;
(e) DVPIVGIITRLTAQKG;
(f) NGQVVLLGSA;
(g)AGSDFIIVPSIFEPCGLTQLVAMRYGS; and
(h)TGGLVDTV

and wherein said wheat starch synthase polypeptide further comprises an amino acid sequence having at least about 85% identity overall to an amino acid sequence set forth in any one of SEQ ID NOS: <400>2, <400>4, <400>6, <400>8 or <400>10.

Preferably, the isolated nucleic acid molecule encodes a starch synthase polypeptide, protein or enzyme that comprises two, more preferably three, more preferably four, more preferably five, more preferably six, more preferably seven and even more preferably eight of the conserved amino acid motifs listed *supra*. Even more preferably, the said amino acid motifs are located in a relative configuration such as that shown for the wheat SSII or wheat SSIII polypeptides listed in Figure 13 herein.

In a preferred embodiment, the isolated nucleic acid molecule encodes a starch synthase polypeptide, protein or enzyme having at least about 90% amino acid sequence identity to any one of SEQ ID NOS:<400>2, <400>4, <400>6, <400>8 or <400>10, more preferably having at least about 95% or about 97% or about 99%

identity to any one of said amino acid sequences.

In an alternative embodiment, the present invention provides an isolated nucleic acid molecule which encodes a wheat starch synthase polypeptide, protein or enzyme molecule or a functional subunit thereof, wherein said nucleic acid molecule comprises a nucleotide sequence having at least about 85% nucleotide sequence identity to any one of SEQ ID NOS: <400>1, <400>3, <400>5, <400>7 or <400>9 or SEQ ID NOS: <400>11 to <400>16, or a degenerate nucleotide sequence thereto or a complementary nucleotide sequence thereto.

By "degenerate nucleotide sequence" is meant a nucleotide sequence that encodes a substantially identical amino acid sequence as a stated nucleotide sequence.

In a preferred embodiment, the isolated nucleic acid molecule comprises the nucleotide sequence set forth in any one of SEQ ID NOS: <400>1, <400>3, <400>5, <400>7 or <400>9 or SEQ ID NOS: <400>11 to <400>16, or is at least about 90% identical, more preferably at least about 95% or 97% or 99% identical to all or a protein-encoding part thereof.

In an alternative embodiment, preferred homologues, analogues and derivatives of the nucleic acid molecule of the present invention encodes a wheat starch synthase polypeptide, protein or enzyme molecule or a functional subunit thereof and comprises a nucleotide sequence that is capable of hybridising under at least moderate stringency hybridisation conditions to at least about 30 contiguous nucleotides derived from any one of SEQ ID NOS: <400>1, <400>3, <400>5, <400>7 or <400>9 or SEQ ID NOS: <400>11 to <400>16, or a complementary nucleotide sequence thereto.

For the purposes of defining the level of stringency, a low stringency is defined herein as being a hybridisation and/or a wash carried out in 6xSSC buffer, 0.1% (w/v) SDS at 28°C. Generally, the stringency is increased by reducing the concentration of SSC

buffer, and/or increasing the concentration of SDS and/or increasing the temperature of the hybridisation and/or wash. A moderate stringency comprises a hybridisation and/or a wash carried out in 0.2 x SSC-2 x SSC buffer, 0.1% (w/v) SDS at 42°C to 65°C, while a high stringency comprises a hybridisation and/or a wash carried out in 0.1xSSC-0.2 x SSC buffer, 0.1% (w/v) SDS at a temperature of at least 55°C. Conditions for hybridisations and washes are well understood by one normally skilled in the art. For the purposes of further clarification only, reference to the parameters affecting hybridisation between nucleic acid molecules is found in pages 2.10.8 to 2.10.16. of Ausubel *et al.* (1987), which is herein incorporated by reference.

Those skilled in the art will be aware of procedures for the isolation of further wheat starch synthase genes to those specifically described herein or homologues, analogues or derivatives of said genes, for example further cDNA sequences and genomic gene equivalents, when provided with one or more of the nucleotide sequences set forth in SEQ ID NOs: <400>1, <400>3, <400>5, <400>7, <400>9, or <400>11 to <400>16. In particular, amplifications and/or hybridisations may be performed using one or more nucleic acid primers or hybridisation probes comprising at least 10 contiguous nucleotides and preferably at least about 20 contiguous nucleotides or 50 contiguous nucleotides derived from the nucleotide sequences set forth herein, to isolate cDNA clones, mRNA molecules, genomic clones from a genomic library (in particular genomic clones containing the entire 5' upstream region of the gene including the promoter sequence, and the entire coding region and 3'-untranslated sequences), and/or synthetic oligonucleotide molecules, amongst others. The present invention clearly extends to such related sequences.

Accordingly, a second aspect of the present invention provides a method of isolating a nucleic acid molecule that encodes a starch synthase polypeptide, protein or enzyme having at least about 85% amino acid sequence identity to any one SEQ ID NOS:<400>2, <400>4, <400>6, <400>8 or <400>10 and/or which comprises a conserved amino acid sequence having at least 25% identity to an amino acid

sequence selected from the group consisting of:

(a) KVGGLGDVVTS;

(b) GHTVEVILPKY;

(c) HDWSSAPVAWLYKEHY;

(d) GILNGIDPDIWDPYTD;

(e) DVPIVGIITRLTAQKG;

(f) NGQVVLLGSA;

(g)AGSDFIIVPSIFEPCGLTQLVAMRYGS; and

(h)TGGLVDTV,

said method comprising:

(i) hybridising a probe or primer comprising at least about 15 contiguous nucleotides in length derived from any one of SEQ ID NOS: <400>1, <400>3, <400>5, <400>7 or <400>9 or SEQ ID NOS: <400>11 to <400>16, or a complementary nucleotide sequence thereto to single-stranded or double-stranded mRNA, cDNA or genomic DNA; and

(ii) detecting the hybridised mRNA, cDNA or genomic DNA using a detecting means.

Preferably, the detecting means is a reporter molecule covalently attached to the probe or primer molecule or alternatively, a polymerase chain reaction format.

An alternative method contemplated in the present invention involves hybridising two nucleic acid "primer molecules" to a nucleic acid "template molecule" which comprises a related starch synthase gene or related starch synthase genetic sequence or a functional part thereof, wherein the first of said primers comprises contiguous nucleotides derived from any one or more of SEQ ID NOS: <400>1, <400>3, <400>5, <400>7 or <400>9 or SEQ ID NOS: <400>11 to <400>16 and the second of said primers comprises contiguous nucleotides complementary to any one or more of SEQ ID NOS: <400>1, <400>3, <400>5, <400>7 or <400>9 or SEQ ID NOS: <400>11 to <400>16. Specific nucleic acid molecule copies of the template molecule are amplified

enzymatically in a polymerase chain reaction, a technique that is well known to one skilled in the art.

In a preferred embodiment, each nucleic acid primer molecule is at least 10 nucleotides in length, more preferably at least 20 nucleotides in length, even more preferably at least 30 nucleotides in length, still more preferably at least 40 nucleotides in length and even still more preferably at least 50 nucleotides in length.

Furthermore, the nucleic acid primer molecules consists of a combination of any of the nucleotides adenine, cytidine, guanine, thymidine, or inosine, or functional analogues or derivatives thereof which are at least capable of being incorporated into a polynucleotide molecule without having an inhibitory effect on the hybridisation of said primer to the template molecule in the environment in which it is used.

Furthermore, one or both of the nucleic acid primer molecules may be contained in an aqueous mixture of other nucleic acid primer molecules, for example a mixture of degenerate primer sequences which vary from each other by one or more nucleotide substitutions or deletions. Alternatively, one or both of the nucleic acid primer molecules may be in a substantially pure form.

The nucleic acid template molecule may be in a recombinant form, in a virus particle, bacteriophage particle, yeast cell, animal cell, or a plant cell. Preferably, the nucleic acid template molecule is derived from a plant cell, tissue or organ, in particular a cell, tissue or organ derived from a wheat or barley plant or a progenitor species, or a relative thereto such as the diploid *Triticum tauschii* or other diploid, tetraploid, aneuploid, polyploid, nullisomic, or a wheat/barley addition line, amongst others.

Those skilled in the art will be aware that there are many known variations of the basic polymerase chain reaction procedure, which may be employed to isolate a related starch synthase gene or related starch synthase genetic sequence when provided with

the nucleotide sequences set forth herein. Such variations are discussed, for example, in McPherson *et al* (1991). The present invention extends to the use of all such variations in the isolation of related starch synthase genes or related starch synthase genetic sequences using the nucleotide sequences embodied by the present invention.

As exemplified herein, the present inventors have isolated several wheat starch synthase genes using both hybridisation and polymerase chain reaction approaches, employing novel probes and primer sequences to do so.

Accordingly, a third aspect of the invention provides an isolated probe or primer comprising at least about 15 contiguous nucleotides in length derived from any one of SEQ ID NOS: <400>1, <400>3, <400>5, <400>7 or <400>9 or SEQ ID NOS: <400>11 to <400>16, or a complementary nucleotide sequence thereto.

Preferably, the probe or primer comprises a nucleotide sequence set forth in any one of SEQ ID NOS:<400>25 to <400>34.

The isolated nucleic acid molecule of the present invention may be introduced into and expressed in any cell, for example a plant cell, fungal cell, insect cell. animal cell, yeast cell or bacterial cell. Those skilled in the art will be aware of any modifications which are required to the codon usage or promoter sequences or other regulatory sequences, in order for expression to occur in such cells.

A further aspect of the invention provides a method of assaying for the presence or absence of a starch synthase isoenzyme or the copy number of a gene encoding same in a plant, comprising contacting a biological sample derived from said plant with an isolated nucleic acid molecule derived from any one of SEQ ID NOS: <400>1, <400>3, <400>5, <400>7 or <400>9 or any one of SEQ ID NOS: <400>11 to <400>16, or any one of SEQ ID NOS:<400>25 to <400>34, or a complementary nucleotide sequence thereto for a time and under conditions sufficient for hybridisation to occur and then

detecting said hybridisation using a detection means.

The detection means according to this aspect of the invention is any nucleic acid based hybridisation or amplification reaction.

The hexaploid nature of wheat prevents the straightforward identification of starch synthase allelic variants by hybridisation using the complete starch synthase-encoding sequence, because the similarities between the various alleles generally results in significant cross-hybridisation. Accordingly, sequence-specific hybridisation probes are required to distinguish between the various alleles. Similarly, wherein PCR is used to amplify specific allelelic variants of a starch synthase gene, one or more sequence-specific amplification primers are generally required. As will be apparent from the amino acid sequence comparisons provided herein, such as in Figures 3 and 13, non-conserved regions of particular wheat starch synthase polypeptides are particularly useful for the design of probes and primers that are capable of distinguishing between one or more starch synthase polypeptide isoenzyme or allelic variant. The present invention clearly contemplates the design of such probes and primers based upon the sequence comparisons provided herein.

In the performance of this embodiment of the present invention, the present inventors particularly contemplate the identification of wheat starch synthase null alleles or alternatively, mutations wherein specific amino acids are inserted or deleted or substituted , compared to one or more of the wheat SSII or SSIII alleles disclosed herein. Such null alleles and other allelic variants are readily identifiable using PCR screeening which employs amplification primers based upon the nucleotide and amino acid sequences disclosed herein for SSII and/or SSIII. Once identified, the various mutations can be stacked or pyramided into one or more new wheat lines, such as by introgression and/or standard plant breeding and/or recombinant approaches (eg. transformation, transfection, etc) thereby producing a novel germplasm which exhibits altered starch properties compared to existing lines. DNA markers based upon the

nucleotide and amino acid sequences disclosed herein for SSII and/or SSIII can be employed to monitor the stacking of genes into the new lines and to correlate the presence of particular genes with starch phenotypes of said lines.

In this regard, a significant advantage conferred by the present invention is the design of new DNA markers that reveal polymorphisms such as, for example, length polymorphisms, restriction site polymorphisms, and single nucleotide polymorphisms, amongst others, between wheat starch synthases and, in particular, between wheat GBSS and/or SSI and/or SSII and/or SSIII, or between allelic variants of one or more of said starch synthases, that can be used to identify the three genomes of hexaploid wheats (i.e., the A, B and D genomes).

Preferably, such DNA markers are derived from the intron region of a starch synthase gene disclosed herein, more preferably the wheat SSII and/or the wheat SSIII gene. Those skilled in the art will be aware that such regions generally have a higher degree of variation than in the protein-encoding regions and, as a consequence, are particularly useful in identifying specific allelic variants of a particular gene, such as allelic variants contained in any one of the three wheat genomes, or alternatively or in addition, for the purpose of distinguishing between wheat GBSS, SSI, SSII or SSIII genes.

A further approach contemplated by the present inventors is the design of unique isoenzyme-specific and/or allele-specific peptides based upon the amino acid sequence disclosed herein as SEQ ID NOS:<400>2 and/or <400>4 and/or <400>6 and/or <400>8 and/or <400>10, which peptides are then used to produce polyclonal or monoclonal antibodies by conventional means. Alternatively, the genes encoding these polypeptides or unique peptide regions thereof can be introduced in an expressible format into an appropriate prokaryotic or eukaroyotic expression system, where they can be expressed to produce the isoenzyme-specific and/or allele-specific peptides for antibody production. Such antibodies may also be used as markers for the

purpose of both identifying parental lines and germplasms and monitoring the stacking of genes in new lines, using conventional immunoassays such as, for example, ELISA and western blotting.

A further aspect of the present invention utilises the above-mentioned nucleic acid based assay method in the breeding and/or selection of plants which express or do not express particular starch synthase isoenzymes or alternatively, which express a particular starch synthase isoenzyme at a particular level in one or more plant tissues. This aspect clearly extends to the selection of transformed plant material which contains one or more of the isolated nucleic acid molecules of the present invention.

Yet another aspect of the present invention provides for the expression of the nucleic acid molecule of the present invention in a suitable host (e.g. a prokaryote or eukaryote) to produce full length or non-full length recombinant starch synthase gene products.

Hereinafter the term "starch synthase gene product" shall be taken to refer to a recombinant product of a starch synthase gene of the present invention.

Preferably, the recombinant starch synthase gene product comprises an amino acid sequence having the catalytic activity of a starch synthase polypeptide or a functional mutant, derivative part, fragment, or analogue thereof.

In a particularly preferred embodiment of the invention, the recombinant starch synthase gene product is selected from the following:

(i) a wheat starch synthase II (wSSII) polypeptide, protein or enzyme or functional subunit thereof which comprises an amino acid sequence which is at least about 85% identical overall to an amino acid sequence set forth in any one of SEQ ID NOS: <400>2, <400>4, or <400>6;

(ii) a wheat starch synthase III (wSSIII) polypeptide, protein or enzyme or

functional subunit thereof which comprises an amino acid sequence which is at least about 85% identical overall to an amino acid sequence set forth in any one of SEQ ID NOS: <400>8 or <400>10;and

(iii) a wheat starch synthase polypeptide, protein or enzyme or functional subunit thereof which comprises a conserved amino acid sequence having at least 25% identity to an amino acid sequence selected from the group consisting of:

(a) KVGGLGDVVTS;
(b) GHTVEVILPKY;
(c) HDWSSAPVAWLYKEHY;
(d) GILNGIDPDIWDPYTD;
(e) DVPIVGIITRLTAQKG;
(f) NGQVVLLGSA;
(g)AGSDFIIVPSIFEPCGLTQLVAMRYGS; and
(h)TGGLVDTV

and which is at least about 85% identical overall to an amino acid sequence set forth in any one of SEQ ID NOS: <400>2, <400>4, <400>6, <400>8 or <400>10.

Accordingly, the present invention clearly extends to homologues, analogues and derivatives of the amino acid sequences set forth herein as SEQ ID NOS: <400>2, <400>4, <400>6, <400>8 and <400>10.

In the present context, "homologues" of an amino acid sequence refer to those polypeptides, enzymes or proteins which have a similar catalytic activity to the amino acid sequences described herein, notwithstanding any amino acid substitutions, additions or deletions thereto. A homologue may be isolated or derived from the same or another plant species as the species from which the polypeptides of the invention are derived.

"Analogues" encompass polypeptides of the invention notwithstanding the occurrence of any non-naturally occurring amino acid analogues therein.

"Derivatives" include modified peptides in which ligands are attached to one or more of the amino acid residues contained therein, such as carbohydrates, enzymes, proteins, polypeptides or reporter molecules such as radionuclides or fluorescent compounds. Glycosylated, fluorescent, acylated or alkylated forms of the subject peptides are particularly contemplated by the present invention. Additionally, derivatives of an amino acid sequence described herein which comprises fragments or parts of the subject amino acid sequences are within the scope of the invention, as are homopolymers or heteropolymers comprising two or more copies of the subject polypeptides. Procedures for derivatizing peptides are well-known in the art.

Substitutions encompass amino acid alterations in which an amino acid is replaced with a different naturally-occurring or a non-conventional amino acid residue. Such substitutions may be classified as "conservative", in which an amino acid residue contained in a starch synthase gene product is replaced with another naturally-occurring amino acid of similar character, for example Gly↔Ala, Val↔Ile↔Leu, Asp↔Glu, Lys↔Arg, Asn↔Gln or Phe↔Trp↔Tyr.

Substitutions encompassed by the present invention may also be "non-conservative", in which an amino acid residue which is present in a starch synthase gene product described herein is substituted with an amino acid with different properties, such as a naturally-occurring amino acid from a different group (eg. substituted a charged or hydrophobic amino acid with alanine), or alternatively, in which a naturally-occurring amino acid is substituted with a non-conventional amino acid.

Non-conventional amino acids encompassed by the invention include, but are not limited to those listed in Table 2.

Amino acid substitutions are typically of single residues, but may be of multiple residues, either clustered or dispersed.

Amino acid deletions will usually be of the order of about 1-10 amino acid residues, while insertions may be of any length. Deletions and insertions may be made to the N-terminus, the C-terminus or be internal deletions or insertions. Generally, insertions within the amino acid sequence will be smaller than amino- or carboxy-terminal fusions and of the order of 1-4 amino acid residues.

A homologue, analogue or derivative of a starch synthase gene product as referred to herein may readily be made using peptide synthetic techniques well-known in the art, such as solid phase peptide synthesis and the like, or by recombinant DNA manipulations. Techniques for making substituent mutations at pre-determined sites using recombinant DNA technology, for example by M13 mutagenesis, are also well-known. The manipulation of nucleic acid molecules to produce variant peptides, polypeptides or proteins which manifest as substitutions, insertions or deletions are well-known in the art.

The starch synthase gene products described herein may be derivatized further by the inclusion or attachment thereto of a protective group which prevents, inhibits or slows proteolytic or cellular degradative processes. Such derivatization may be useful where the half-life of the subject polypeptide is required to be extended, for example to increase the amount of starch produced in the endosperm or alternatively, to increase the amount of protein produced in a bacterial or eukaryotic expression system. Examples of chemical groups suitable for this purpose include, but are not limited to, any of the non-conventional amino acid residues listed in Table 2, in particular a D-stereoisomer or a methylated form of a naturally-occurring amino acid listed in Table 1. Additional chemical groups which are useful for this purpose are selected from the list comprising aryl or heterocyclic N-acyl substituents, polyalkylene oxide moieties, desulphatohirudin muteins, alpha-muteins, alpha-aminophosphonic acids, water-

soluble polymer groups such as polyethylene glycol attached to sugar residues using hydrazone or oxime groups, benzodiazepine dione derivatives, glycosyl groups such as beta-glycosylamine or a derivative thereof, isocyanate conjugated to a polyol functional group or polyoxyethylene polyol capped with diisocyanate, amongst others. Similarly, a starch synthase gene product or a homologue, analogue or derivative thereof may be cross-linked or fused to itself or to a protease inhibitor peptide, to reduce susceptibility of said molecule to proteolysis.

In a particularly preferred embodiment, the percentage similarity to in any one of SEQ ID NOS: <400>2, <400>4, <400>6, <400>8 or <400>10 is at least about 90%, more preferably at least about 95%, even more preferably at least about 97% and even more preferably at least about 98%, or about 99% or 100%.

In a related embodiment, the present invention provides a "sequencably pure" form of the amino acid sequence described herein. "Sequencably pure" is hereinbefore described as substantially homogeneous to facilitate amino acid determination.

In a further related embodiment, the present invention provides a "substantially homogeneous" form of the subject amino acid sequence, wherein the term "substantially homogeneous" is hereinbefore defined as being in a form suitable for interaction with an immunologically interactive molecule. Preferably, the polypeptide is at least 20% homogeneous, more preferably at least 50% homogeneous, still more preferably at least 75% homogeneous and yet still more preferably at least about 95-100% homogenous, in terms of activity per microgram of total protein in the protein preparation.

To produce the recombinant polypeptide of the present invention, the coding region of a starch synthase gene described herein or a functional homologue, analogue or derivative thereof is placed operably in connection with a promoter sequence in the sense orientation, such that a starch synthase gene product is capable of being

expressed under the control of said promoter sequence.

In the present context, the term "in operable connection with" means that expression of the isolated nucleotide sequence is under the control of the promoter sequence with which it is connected, regardless of the relative physical distance of the sequences from each other or their relative orientation with respect to each other.

Reference herein to a "promoter" is to be taken in its broadest context and includes the transcriptional regulatory sequences of a classical genomic gene, including the TATA box which is required for accurate transcription initiation, with or without a CCAAT box sequence and additional regulatory elements (i.e. upstream activating sequences, enhancers and silencers) which alter gene expression in response to developmental and/or external stimuli, or in a tissue-specific manner. A promoter is usually, but not necessarily, positioned upstream or 5′, of a structural gene, the expression of which it regulates. Furthermore, the regulatory elements comprising a promoter are usually positioned within 2 kb of the start site of transcription of the gene.

In the present context, the term "promoter" is also used to describe a synthetic or fusion molecule, or derivative which confers, activates or enhances expression of a structural gene or other nucleic acid molecule, particularly in a plant cell and more preferably in a wheat plant or other monocotyledonous plant cell, tissue or organ. Preferred promoters may contain additional copies of one or more specific regulatory elements, to further enhance expression and/or to alter the spatial expression and/or temporal expression. For example, regulatory elements which confer copper inducibility may be placed adjacent to a heterologous promoter sequence, thereby conferring copper inducibility on the expression of said molecule.

Those skilled in the art will be aware that in order to obtain optimum expression of the starch synthase gene of the present invention, it is necessary to position said gene in an appropriate configuration such that expression is controlled by the promoter

sequence. Promoters are generally positioned 5′ (upstream) to the genes that they control. In the construction of heterologous promoter/structural gene combinations it is generally preferred to position the promoter at a distance from the gene transcription start site that is approximately the same as the distance between that promoter and the gene it controls in its natural setting, i.e., the gene from which the promoter is derived. As is known in the art, some variation in this distance can be accommodated without loss of promoter function. Similarly, the preferred positioning of a regulatory sequence element with respect to a heterologous gene to be placed under its control is defined by the positioning of the element in its natural setting, i.e., the genes from which it is derived. Again, as is known in the art, some variation in this distance can also occur.

Examples of promoters suitable for expressing the starch synthase gene of the present invention include viral, fungal, bacterial, animal and plant derived promoters capable of functioning in prokaryotic or eukaryotic cells. Preferred promoters are those capable of regulating the expression of the subject starch synthase genes in plants cells, fungal cells, insect cells, yeast cells, animal cells or bacterial cells, amongst others. Particularly preferred promoters are capable of regulating expression of the subject nucleic acid molecules in monocotyledonous plant cells. The promoter may regulate the expression of the said molecule constitutively, or differentially with respect to the tissue in which expression occurs or, with respect to the developmental stage at which expression occurs, or in response to external stimuli such as physiological stresses, or plant pathogens, or metal ions, amongst others.

Accordingly, strong constitutive promoters are particularly preferred for the purposes of the present invention.

Examples of preferred promoters include the bacteriophage T7 promoter, bacteriophage T3 promoter, SP6 promoter, *lac* operator-promoter, *tac* promoter, SV40 late promoter, SV40 early promoter, RSV-LTR promoter, CMV IE promoter, CaMV 35S

promoter, SCSV promoter, SCBV promoter and the like.

Particularly preferred promoters operable in plant cells include, for example the CaMV 35S promoter, and the SCBV promoter. Those skilled in the art will readily be aware of additional promoter sequences other than those specifically described.

In a particularly preferred embodiment, the promoter may be derived from a genomic starch synthase gene. Preferably, the promoter sequence comprises nucleotide sequences that are linked *in vivo* to nucleotide sequences set forth in any one of SEQ ID NOs: <400>1, <400>3, <400>5, <400>7, <400>9, or any one of SEQ ID NOs: <400>11 to <400>16. By "linked *in vivo*" means that the promoter is present in its native state in the genome of a wheat plant where it controls expression of the starch synthase gene of the present invention.

Conveniently, genetic constructs are employed to facilitate expression of a starch synthase genetic sequence of the present invention or a functional derivative, part, homologue, or analogue thereof. To produce a genetic construct, the starch synthase gene of the invention is inserted into a suitable vector or episome molecule, such as a bacteriophage vector, viral vector or a plasmid, cosmid or artificial chromosome vector which is capable of being maintained and/or replicated and/or expressed in the host cell, tissue or organ into which it is subsequently introduced. The said genetic construct comprises the subject nucleic acid molecule placed operably under the control of a promoter sequence and optionally, a terminator sequence.

The term "terminator" refers to a DNA sequence at the end of a transcriptional unit which signals termination of transcription. Terminators are 3′-non-translated DNA sequences containing a polyadenylation signal, which facilitates the addition of polyadenylate sequences to the 3′-end of a primary transcript. Terminators active in bacteria, yeasts, animal cells and plant cells are known and described in the literature. They may be isolated from bacteria, fungi, viruses, animals and/or plants.

Examples of terminators particularly suitable for use in expressing the nucleic acid molecule of the present invention in plant cells include the nopaline synthase (NOS) gene terminator of *Agrobacterium tumefaciens,* the terminator of the Cauliflower mosaic virus (CaMV) 35S gene, and the *zein* gene terminator from *Zea mays.*

Genetic constructs will generally further comprise one or more origins of replication and/or selectable marker gene sequences.

The origin of replication can be functional in a bacterial cell and comprise, for example, the pUC or the ColE1 origin. Alternatively, the origin of replication is operable in a eukaryotic cell, tissue and more preferably comprises the 2 micron ($2\mu m$) origin of replication or the SV40 origin of replication.

As used herein, the term "selectable marker gene" includes any gene which confers a phenotype on a cell in which it is expressed to facilitate the identification and/or selection of cells which are transfected or transformed with a genetic construct of the invention or a derivative thereof.

Suitable selectable marker genes contemplated herein include the ampicillin-resistance gene (Amp^r), tetracycline-resistance gene (Tc^r), bacterial kanamycin-resistance gene (Kan^r), is the zeocin resistance gene (Zeocin is a drug of bleomycin family which is trademark of InVitrogen Corporation), the *AURI-C* gene which confers resistance to the antibiotic aureobasidin A, phosphinothricin-resistance gene, neomycin phosphotransferase gene (*nptII*), hygromycin-resistance gene, β-glucuronidase (GUS) gene, chloramphenicol acetyltransferase (CAT) gene, green fluorescent protein-encoding gene or the luciferase gene, amongst others. Those skilled in the art will be aware of other selectable marker genes useful in the performance of the present invention and the subject invention is not limited by the nature of the selectable marker gene.

Usually, an origin of replication or a selectable marker gene suitable for use in bacteria is physically-separated from those genetic sequences contained in the genetic construct which are intended to be expressed or transferred to a eukaryotic cell, or integrated into the genome of a eukaryotic cell.

Standard methods can be used to introduce genetic constructs into a cell, tissue or organ for the purposes of modulating gene expression. Particularly preferred methods suited to the introduction of synthetic genes and genetic constructs comprising same to eukaryotic cells include liposome-mediated transfection or transformation, transformation of cells with attenuated virus particles or bacterial cells and standard procedures for the transformation of plant and animal cells, tissues, organs or organisms. Any standard means may be used for their introduction including cell mating, transformation or transfection procedures known to those skilled in the art or described by Ausubel *et al.* (1992).

In a further embodiment of the present invention, the starch synthase genes of the present invention and genetic constructs comprising same are adapted for integration into the genome of a cell in which it is expressed. Those skilled in the art will be aware that, in order to achieve integration of a genetic sequence or genetic construct into the genome of a host cell, certain additional genetic sequences may be required. In the case of plants, left and right border sequences from the T-DNA of the *Agrobacterium tumefaciens* Ti plasmid will generally be required.

The invention further contemplates increased starch and/or modified starch composition in transgenic plants expressing the nucleic acid molecule of the invention in the sense orientation such that the activity of one or more starch synthase isoenzymes is increased therein. By increasing the level of one or more starch synthase isoenzymes, the deposition of starch in the amyloplast or chloroplast is increased and/or a modified starch granule structure is produced and/or starch composition is modified and/or the amylose/amylopectin ratio is altered in the plant.

Wherein it is desired to increase the synthesis of a particular starch synthase isoenzyme in a plant cell, the coding region of a starch synthase gene is placed operably behind a promoter, in the sense orientation, such that said starch synthase is expressed under the control of said promoter sequence. In a preferred embodiment, the starch synthase genetic sequence is a starch synthase genomic sequence, cDNA molecule or protein-coding sequence.

Wherein it is desirable to reduce the level of a particular starch synthase isoenzyme in a plant cell, the nucleic acid molecule of the present invention can be expressed in the antisense orientation, as an antisense molecule or a ribozyme molecule, under the control of a suitable promoter.

Alternatively, the nucleic acid molecule of the present invention may also be expressed in the sense orientation, in the form of a co-suppression molecule, to reduce the level of a particular starch synthase isoenzyme in a plant cell. As will be known to those skilled in the art, co-suppression molecules that comprise inverted repeat sequences of a target nucleic acid molecule provide optimum efficiency at reducing expression of said target nucleic acid molecule and, as a consequence, the present invention clearly contemplates the use of inverted repeat sequences of any one or more of the starch synthase genetic sequences exemplified herein, or inverted repeat sequences of a homologue, analogue or derivative of said starch synthase genetic sequences, to reduce the level of a starch synthase isoenzyme in a plant.

The expression of an antisense, ribozyme or co-suppression molecule comprising a starch synthase gene in a cell such as a plant cell, fungal cell, insect cell. animal cell, yeast cell or bacterial cell, may also increase the availability of carbon as a precursor for a secondary metabolite other than starch (e.g. sucrose or cellulose). By targeting the endogenous starch synthase gene, expression is diminished, reduced or otherwise lowered to a level that results in reduced deposition of starch in the amyloplast or chloroplast and/or leads to modified starch granule structure and/or composition

and/or altered amylose/amylopectin ratio.

Accordingly, a further aspect of the present invention provides a method of modifying the starch content and/or starch composition of one or more tissues or organs of a plant, comprising expressing therein a sense molecule, antisense molecule, ribozyme molecule, co-suppression molecule, or gene-targeting molecule having at least about 85% nucleotide sequence identity to any one of any one of SEQ ID NOS: <400>1, <400>3, <400>5, <400>7 or <400>9 or any one of SEQ ID NOS: <400>11 to <400>16, or a complementary nucleotide sequence thereto for a time and under conditions sufficient for the enzyme activity of one or more starch synthase isoenzymes to be modified. This aspect of the invention clearly extends to the introduction of the sense molecule, antisense molecule, ribozyme molecule, co-suppression molecule, or gene-targeting molecule to isolated plant cells, tissues or organs or organelles by cell fusion or transgenic means and the regeneration of intact plants therefrom.

Co-suppression is the reduction in expression of an endogenous gene that occurs when one or more copies of said gene, or one or more copies of a substantially similar gene are introduced into the cell, preferably in the form of an inverted repeat structure.

The present inventors have discovered that the genetic sequences disclosed herein are capable of being used to modify the level of starch when expressed, particularly when expressed in plants cells. Accordingly, the present invention clearly extends to the modification of starch biosynthesis in plants, in particular wheat or barley plants or a progenitor plant species, or a relative thereto such as the diploid *Triticum tauschii* or other diploid, tetraploid, aneuploid, polyploid, nullisomic, or a wheat/barley addition line, amongst others.

In particular, the present invention contemplates decreased starch production and/or modified starch composition in transgenic plants expressing the nucleic acid molecule of the invention in the antisense orientation or alteratively, expressing a ribozyme or

co-suppression molecule comprising the nucleic acid sequence of the invention such that the activity of one or more starch synthase isoenzymes is decreased therein.

In the context of the present invention, an antisense molecule is an RNA molecule which is transcribed from the complementary strand of a nuclear gene to that which is normally transcribed to produce a "sense" mRNA molecule capable of being translated into a starch synthase polypeptide. The antisense molecule is therefore complementary to the mRNA transcribed from a sense starch synthase gene or a part thereof. Although not limiting the mode of action of the antisense molecules of the present invention to any specific mechanism, the antisense RNA molecule possesses the capacity to form a double-stranded mRNA by base pairing with the sense mRNA, which may prevent translation of the sense mRNA and subsequent synthesis of a polypeptide gene product.

Ribozymes are synthetic RNA molecules which comprise a hybridising region complementary to two regions, each of at least 5 contiguous nucleotide bases in the target sense mRNA. In addition, ribozymes possess highly specific endoribonuclease activity, which autocatalytically cleaves the target sense mRNA. A complete description of the function of ribozymes is presented by Haseloff and Gerlach (1988) and contained in International Patent Application No. WO89/05852.

The present invention extends to ribozyme which target a sense mRNA encoding a native starch synthase gene product, thereby hybridising to said sense mRNA and cleaving it, such that it is no longer capable of being translated to synthesise a functional polypeptide product.

According to this embodiment, the present invention provides a ribozyme or antisense molecule comprising at least 5 contiguous nucleotide bases derived from any one of

SEQ ID NOS: <400>1, <400>3, <400>5, <400>7 or <400>9 or any one of SEQ ID NOS: <400>11 to <400>16, or a complementary nucleotide sequence thereto or a homologue, analogue or derivative thereof, wherein said antisense or ribozyme molecule is able to form a hydrogen-bonded complex with a sense mRNA encoding a starch synthase gene product to reduce translation thereof.

In a preferred embodiment, the antisense or ribozyme molecule comprises at least 10 to 20 contiguous nucleotides derived from any one of SEQ ID NOS: <400>1, <400>3, <400>5, <400>7 or <400>9 or any one of SEQ ID NOS: <400>11 to <400>16, or a complementary nucleotide sequence thereto or a homologue, analogue or derivative thereof. Although the preferred antisense and/or ribozyme molecules hybridise to at least about 10 to 20 nucleotides of the target molecule, the present invention extends to molecules capable of hybridising to at least about 50-100 nucleotide bases in length, or a molecule capable of hybridising to a full-length or substantially full-length mRNA encoded by a starch synthase gene.

Those skilled in the art will be aware of the necessary conditions, if any, for selecting or preparing the antisense or ribozyme molecules of the invention.

It is understood in the art that certain modifications, including nucleotide substitutions amongst others, may be made to the antisense and/or ribozyme molecules of the present invention, without destroying the efficacy of said molecules in inhibiting the expression of a starch synthase gene. It is therefore within the scope of the present invention to include any nucleotide sequence variants, homologues, analogues, or fragments of the said gene encoding same, the only requirement being that said nucleotide sequence variant, when transcribed, produces an antisense and/or ribozyme molecule which is capable of hybridising to a sense mRNA molecule which encodes a starch synthase gene product.

Gene targeting is the replacement of an endogenous gene sequence within a cell by

a related DNA sequence to which it hybridises, thereby altering the form and/or function of the endogenous gene and the subsequent phenotype of the cell. According to this embodiment, at least a part of the DNA sequence defined by any one of SEQ ID NOS: <400>1, <400>3, <400>5, <400>7 or <400>9 or any one of SEQ ID NOS: <400>11 to <400>16 may be introduced into target cells containing an endogenous gene that encodes a particular starch synthase isoenzyme, thereby replacing said endogenous gene. According to this embodiment, the polypeptide product of the gene targetting molecule generally encodes a starch synthase isoenzyme that possesses different catalytic activity to the polypeptide product of the endogenous gene, producing in turn modified starch content and/or composition in the target cell.

The present invention extends to genetic constructs designed to facilitate expression of a sense molecule, an antisense molecule, ribozyme molecule, co-suppression molecule, or gene targeting molecule of the present invention. The requirements for expressing such molecules are similar to those for expressing a recombinant polypeptide as described *supra*.

The present invention further extends to the production and use of starches produced by the application of the novel genes described herein.

Starch hydrolysates or undegraded starches are both useful in industry and, as a consequence, the present invention is useful in applications relating to the use of both starch hydrolysates and undegraded starches. By "starch hydrolysates" is meant the glucose and glucan components that are obtainable by the enzymatic or chemical degradation of starch in chemical modifications and processes, such as fermentation.

For example, starch produced by plants expressing the sense, antisense, co-suppression, gene-targetting or ribozyme molecules of the present invention may exhibit modified viscosities and/or gelling properties of its glues when compared to starch derived from wild-type plants. Native starches produced by the performance of

the inventive method are useful as an additive in the following: (i) foodstuffs, for the purpose of increasing the viscosity or gelling properties of food; (ii) in non-foodstuffs, such as an adjuvant or additive in the paper and cardboard industries, for retention or as a size filler, or as a solidifying substance or for dehydration, or film coating, amongst others; (iii) in the adhesive industry as pure starch glue, as an additive to synthetic resins and polymer dispersions, or as an extenders for synthetic adhesives; (iv) in the textile and textile care industries to strengthen woven products and reduce burring or to thicken dye pastes; (v) in the building industry, such as a binding agent in the production of gypsum plaster boards, or for the deceleration of the sizing process; (vi) in ground stabilization or for the temporary protection of ground particles against water in artificial earth shifting; (vii) as a wetting agent in plant protectants and fertilizers; (viii) as a binding agent in drugs, pharmaceuticals and medicated foodstuff such as vitamins, etc; (ix) as an additive in coal and briquettes; (xi) as a flocculent in the processing of coal ore and slurries; (xii) as a binding agent in casting processes to increase flow resistance and improve binding strength; and (xiii) to improve the technical and optical quality of rubber and plastic products. Additional applications are not excluded.

A further aspect of the present invention provides an isolated promoter that is operable in the endosperm of a monocotyledonous plant cell, tissue or organ, and preferably in the endosperm of a monocotyledonous plant cell, tissue or organ. According to this embodiment, it is preferred that the promoter is derived from a starch synthase gene of the present invention, such as a promoter that is linked *in vivo* to any one of SEQ ID NOS: <400>1, <400>3, <400>5, <400>7 or <400>9 or any one of SEQ ID NOS: <400>11 to <400>16, or a complementary nucleotide sequence thereto.

In a particularly preferred embodiment, the promoter comprises a nucleotide sequence derivable from the 5′-upstream region of SEQ ID NO:<400>11 or a complementary nucleotide sequence thereto, an more preferably comprises nucleotides 1 to about 287 of SEQ ID NO:<400>11 or nucleotides 1 to about 287 of SEQ ID NO:<400>11 or a

complementary nucleotide sequence thereto. The present invention clearly extends to promoter sequences that comprise further nucleotide sequences in the region upstream of the stated nucleotide sequence that are linked *in vivo* to said nucleotide sequence in the wheat genome.

In a related embodiment, the promoter sequence of the present invention will further comprise an exon sequence derived from a starch synthase gene, for example nucleotides 260 to 385 of SEQ ID NO:<400>11 or a complementary nucleotide sequence thereto. Those skilled in the art will be aware that the inclusion of such nucleotide sequences may increase the expression of a heterologous structural gene, the expression of which is controlled thereby.

The present invention further extends to the expression of any structural gene operably under the control of the starch synthase promoter sequence exemplified herein or a functional homologue, analogue or derivative of said promoter sequence.

As with other embodiments described herein for expression in cells, a genetic construct may be employed to effect said expression and the present invention clearly extends to said genetic constructs.

The polypeptide encoded by the structural gene component may be a reporter molecule which is encoded by a gene such as the bacterial β-glucuronidase gene or chloramphenicol acetyltransferase gene or alternatively, the firefly luciferase gene. Alternatively, wherein it is desirable to alter carbon partitioning within the endosperm, the polypeptide may be an enzyme of the starch sucrose biosynthetic pathways. Preferably, the promoter sequence is used to regulate the expression of one or more of the starch synthase genes of the present invention or a sense, antisense, ribozyme, co-suppression or gene-targetting molecule comprising or derived from same.

Recombinant DNA molecules carrying the aforesaid nucleic acid molecule of the

present invention or a sense, antisense, ribozyme, gene-targetting or co-suppression molecule and/or genetic construct comprising same, may be introduced into plant tissue, thereby producing a "transgenic plant", by various techniques known to those skilled in the art. The technique used for a given plant species or specific type of plant tissue depends on the known successful techniques. Means for introducing recombinant DNA into plant tissue include, but are not limited to, transformation (Paszkowski *et al.,* 1984), electroporation (Fromm *et al.,* 1985), or microinjection of the DNA (Crossway *et al.*, 1986), or T-DNA-mediated transfer from *Agrobacterium* to the plant tissue. Representative T-DNA vector systems are described in the following references: An *et al.*(1985); Herrera-Estrella *et al.* (1983a,b); Herrera-Estrella *et al.* (1985). Once introduced into the plant tissue, the expression of the introduced gene may be assayed in a transient expression system, or it may be determined after selection for stable integration within the plant genome. Techniques are known for the *in vitro* culture of plant tissue, and in a number of cases, for regeneration into whole plants. Procedures for transferring the introduced gene from the originally transformed plant into commercially useful cultivars are known to those skilled in the art.

In general, plants are regenerated from transformed plant cells or tissues or organs on hormone-containing media and the regenerated plants may take a variety of forms, such as chimeras of transformed cells and non-transformed cells; clonal transformants (e.g., all cells transformed to contain the expression cassette); grafts of transformed and untransformed tissues (e.g., a transformed root stock grafted to an untransformed scion in citrus species). Transformed plants may be propagated by a variety of means, such as by clonal propagation or classical breeding techniques. For example, a first generation (or T1) transformed plants may be selfed to give homozygous second generation (or T2) transformed plants, and the T2 plants further propagated through classical breeding techniques.

Accordingly, a still further aspect of the present invention contemplates a transgenic plant comprising an introduced sense molecule, antisense molecule, ribozyme

molecule, co-suppression molecule, or gene-targeting molecule having at least about 85% nucleotide sequence identity to any one of any one of SEQ ID NOS: <400>1, <400>3, <400>5, <400>7 or <400>9 or any one of SEQ ID NOS: <400>11 to <400>16, or a complementary nucleotide sequence thereto or a genetic construct comprising same.

The present invention further extends to those plant parts, propagules and progeny of said transgenic plant or derived therefrom, the only requirement being that said propagules and progeny also carry the introduced nucleic acid molecule(s).

The present invention is further described by reference to the following non-limiting examples.

EXAMPLE 1

Plant material

Genetic stocks of hexaploid bread wheat *Triticum aestivum* cv. Chinese Spring with various chromosome additions and deletions were kindly supplied by Dr E. Lagudah (CSIRO Plant Industry, Canberra) and derived from stocks described in Sears and Miller (1985). The hexaploid (*Triticum aestivum*) wheats cv Gabo and cv Wyuna were grown in controlled growth cabinet conditions (18 °C day and 13° C night, with a photoperiod of 16 h). Wheat leaves and florets prior to anthesis, and endosperm were collected over the grain filling period, immediately frozen in liquid nitrogen and stored at -80°C until required.

EXAMPLE 2

Gel Electrophoresis, Antibodies and Immunoblotting

Monoclonal antibodies against the Sgp-1 proteins, and their use in the immunoblotting of SDS-PAGE gels have been described previously (Rahman *et al.*, 1995).

EXAMPLE 3

Preparation of total RNA from wheat

Total RNA was isolated from the leaf, floret and endosperm tissues of wheat essentially as described by Higgins *et al.* (1976) or Rahman *et al.* (1998). RNA was quantified by UV absorption and by separation in 1.4% (w/v) agarose-formaldehyde gels which were then visualised under UV light after staining with ethidium bromide.

EXAMPLE 4

Construction and screening of cDNA libraries

A first cDNA library, an expression cDNA library of wheat endosperm, was constructed from mRNA isolated from wheat cv Chinese Spring. RNA from 5, 7, 9, 11 and 13 days after anthesis was pooled and random primers were used for the first strand of cDNA synthesis. Monoclonal antibodies against 100 -105 kDa proteins in wheat starch granules (Rahman *et al.*, 1995) were used for immunoscreening of the expression cDNA library.

A second cDNA library was constructed from the endosperm mRNA of the hexaploid *Triticum aestivum* cultivar Wyuna, 8 - 12 days after anthesis, as described by Rahman *et al.* (1997). This library was screened with a 85-bp cDNA fragment, wSSIIp1, which was obtained by immunoscreening of the expression cDNA library as described above. The wSSIIp1 probe corresponded to nucleotide positions 988 to 1072 of wSSIIB (SEQ ID NO:<400>1) at the hybridisation conditions as described earlier (Rahman *et al.*, 1998).

A third cDNA library was constructed from RNA from the endosperm of the hexaploid *Triticum aestivum* cultivar Rosella as described by Rahman *et al.* (1997). This library was screened with a 347-bp cDNA fragment, wSSIIIp1 for the first screening and a 478-bp cDNA fragment wSSIIIp3 for the second screening (**PLEASE ADVISE-** nucleotides 2469 to 2947 of SEQ ID NO:<400>7) using the hybridisation conditions described herein.

EXAMPLE 5

Construction and screening of *Triticum tauschii* genomic library

The genomic library used in this study, prepared from *Triticum tauschii*, var strangulata, (Accession Number CPI 110799), has been described in Rahman *et al.*, (1997). Of all the accessions of *T. tauschii* surveyed, DNA marker analysis suggests that the genome of CPI 110799 is the most closely related to the D genome of hexaploid wheat (Lagudah *et al.*, 1991).

Hybridisations were carried out in 25% formamide, 6 x SSC, 0.1% SDS at 42°C for 16 hours, then filters were washed 3 times using 2 x SSC containing 0.1% SDS at 65°C for 1 hour per wash.

For the isolation of a genomic wSSII clone, the probe comprised the PCR-derived DNA fragment wSSIIp2 and positive-hybridising plaques were digested using the restriction enzyme *Bam*HI, separated on a 1% agarose gel, transferred to nitrocellulose membrane and hybridised to probe wSSIIp4 comprising nucleotides 1 to 367 of the wSSIIA cDNA clone, using the conditions described by Rahman *et al.* (1997).

For the isolation of a genomic wSSIII clone, plaques hybridising to the PCR-derived DNA fragment wSSIIIp1 from clone wSSIII.B3 (i.e. nucleotides 3620 to 3966 of SEQ ID NO:<400>7) were selected and re-screened until plaque-purified.

EXAMPLE 6

DNA sequencing and analysis

DNA sequencing was performed using the automated ABI system with dye terminators as described by the manufacturers. DNA sequences were analysed using the GCG suite of programs (Devereaux *et al.*, 1984).

EXAMPLE 7

DNA and RNA analysis

DNA was isolated and analysed as previously described (Maniatis *et al.*, 1982; Rahman *et al.*, 1998). Approximately 20 μg of DNA was digested with restriction enzymes *Bam*HI, *Dra*I and *Eco*RI, separated on a 1% agarose gel and transferred to reinforced nitrocellulose membranes (BioRad) and hybridised with ^{32}P-labelled DNA probe, either wSSIIIp1, corresponding to nucleotides 3620 to 3966 of the wheat SSIII gene, or alternatively, with the entire wSSII cDNA clone. DNA fragment probes were labelled with the Rapid Multiprime DNA Probe Labelling Kit (Promega).

The hybridisation and wash conditions were performed as described in Rahman *et al.* (1997). For RNA analysis, 10 μg of total RNA was separated in a 1.4% agarose-formaldehyde gel and transferred to a Hybond N+ membrane (Amersham), and hybridised with cDNA probe at 42°C as previously described by Khandjian *et al.*, (1987) or Rahman *et al.*, (1998). After washing for 30 minutes at 65°C with 2x SSC, 0.1% SDS; followed by three washes of 40 minutes at 65°C with 0.2x SSC, 1% SDS, the membranes were visualised by overnight exposure at -80°C with Kodak MR X-ray film.

EXAMPLE 8

Expression of wheat Sgp-1 polypeptides in the wheat endosperm

The development and use of monoclonal antibodies to the Sgp-1 proteins has been described previously (Rahman *et al.*, 1995). These antibodies were used by the present inventors to characterise the expression and localisation of the Sgp-1 proteins.

The proteins found in the matrix of the wheat starch granule are shown in Figure 1, lane1. The remaining lanes show an immunoblot of proteins from the soluble phase (Figure 1; lanes 2-4) and the starch granule (Figure 1; lanes 5-7), respectively, following SDS-PAGE. In addition to cross-reactivity with the 100-105 kDa proteins, a

weak cross-reaction with a 50 kDa protein in both the granule and the soluble fractions were observed (Figure 1). The Sgp-1 polypeptides are present in the starch granule throughout endosperm development (Figure 1; lanes 5-7, also see Rahman *et al.*, 1995). However, as the endosperms matures, there is a reduction in the amount of Sgp-1 protein found in the soluble fraction. Lane 4 shows that by 25 days after anthesis, the level of these proteins in the soluble fraction is substantially reduced. This observation is consistent with previous results from Rahman *et al.*, (1995), who suggested that the Sgp-1 proteins were exclusively granule bound based on studies of granules from endosperm in mid-late stages endosperm development, however, these results suggest that the partitioning of these proteins between the granule and the soluble phase changes during development.

EXAMPLE 9

Isolation of cDNA clones encoding wheat starch synthase II (wSSII) proteins

Monoclonal antibodies against Sgp-1 polypeptides (Rahman *et al.*, 1995) were used to probe the expression library described in Example 4 (i.e. the first cDNA library). Three immunoreactive plaques were identified and sequenced. One clone, designated wSSIIp1, contained an 85-bp cDNA insert with homology to maize SSIIa (Harn *et al.*, 1998).

DNA from the wSSIIp1 clone was used as a probe in the hybridisation screening of the second cDNA library, prepared from *Triticum aestivum* cultivar Wyuna endosperm RNA as described in Example 4. Ten hybridising cDNA clones were selected and sequenced. On the basis of the DNA sequences obtained, the 10 cDNA clones can be classified into three groups. Group 1 contains 7 cDNA clones, group 2 contains 2 cDNA clones and group 3 contains 1 cDNA clone.

The longest clone from group 1 (designated wSSIIB) is 2939 bp in length (SEQ ID NO:<400>1) and encodes a 798 -amino acid polypeptide starting at nucleotide 176 and terminating at nucleotide 2572 (SEQ ID NO:<400>2).

The longest clone from group 2 (designated wSSIIA) is 2807 bp in length (SEQ ID NO:<400>3) and encodes a 799 -amino acid polypeptide starting at nucleotide 89 and terminating at nucleotide 2488 (SEQ ID NO:<400>4).

The cDNA from group 3 is a partial cDNA clone (designated wSSIID), which is 2107 bp in length (SEQ ID NO:<400>5) and encodes a 597 -amino acid polypeptide starting at nucleotide 1 and terminating at nucleotide 1794 (SEQ ID NO:<400>6). The encoded polypeptide is approximately a 200 amino acid residues shorter than that of polypeptides encoded by longest clones of group 1 or 2 clones, respectively (Figure 2).

Comparison of the three cDNA clones, wSSIIB, wSSIIA and wSSIID shows that they share 95.7% to 96.6% identity at amino acid level, with variation at 44 amino acid positions between the three sequences (Figure 3). Of the 44 amino acid changes between these sequences, 31 changes occur in the N-terminal region (residues 1 to 300), 10 changes occur in the central region (residues 301 to 729) and 3 changes occur in the C-terminal region (residues 730 to 799). The wSSIIA polypeptide (799 amino acid residues) and wSSIIB polypeptide (798 amino acid residues) sequences differ in length by a single amino acid residue, due to the deletion of Asp-69 from the wSSIIB polypeptidesequence.

A comparison of the nucleotide sequences of the wSSIA, wSSIIB and wSSIID cDNA clones with the nucleotide sequence of the wSSIIp1 cDNA obtained by immunoscreening confirms that the wSSIIp1 sequence is found in each cDNA. The peptide encoded by the wSSIIp1 cDNA clone corresponds to amino acid residues in the region from residue 272 to residue 298 of the wSSIIA polypeptide, and to amino acid residues in the region from residue 271 to residue 297 of the wSSIIB polypeptide (see Figure 3). Thus, the peptide epitope encoded by wSSIIp1 that reacts with the anti-Sgp-1 monoclonal antibodies can therefore be localised to this region of the wSSIIA and wSSIIB polypeptides and to the corresponding region of the wSSIID

polypeptide.

Notwithstanding that a region having about 63% amino acid sequence identity to the peptide epitope encoded by clone wSSIIp1 is found in the maize SSIIa polypeptide (Figure 3), the degree of amino acid conservation between maize and wheat sequences in this region of the polypeptide is insufficient for immunological cross-reactivity to occur between these species using the monoclonal antibodies to the wheat Sgp-1 proteins described by Rahman *et al.* (1995). Additionally, this peptide epitope is not found in granule-bound starch synthases, SSI, or SSIII (data not shown).

The wSSIIB cDNA (SEQ ID NO:<400>1) encodes an amino acid sequence comprising the peptide motif AAGKKDAGID (SEQ ID NO:<400>18) between residues 60 and 69 of SEQ ID NO:<400>2 (Figure 3) which, with the exception of the second residue, is identical to the N-terminal of the 100 kDa ($A^{T}/_{L}$GKKDAGID: SEQ ID NOs:<400>19 and 20) protein (Sgp-B1) from the wheat starch granule (note that the sequence given in Rahman *et al.*, 1995 ($A^{T}/_{L}$GKKDAL: SEQ ID NOs:<400>21 and 22) has been revised following further amino acid sequence analysis).

The wSSIIA cDNA clone (SEQ ID NO:<400>3) encodes an amino acid sequence comprising the peptide motif AAGKKDARVDDAA (SEQ ID NO: <400>23) at residues 60 to 73 of SEQ ID NO:<400>4, which is about 66% identical to the N-terminal amino acid sequence (i.e. ALGKKDAGIVDGA: SEQ ID NO: <400>24) of the 104 kDa and 105 kDa starch granule proteins, Sgp-D1 and Sgp-A1 respectively, as determined by sequence analysis of isolated protein (Rahman *et al.*, 1995).

Furthermore, Takaoka *et al.* (1997) reported the amino acid sequences of 3 polypeptides obtained from sequencing starch granule proteins derived from the Sgp-1 proteins. Peptide 3 described by Takaoka *et al.* (1997) corresponds to amino acid residues 378 to 387 of the amino acid sequence of the wSSIIA cDNA (SEQ ID NO:<400>4; Figure 3). Peptides 1 and 2 described by Takaoka *et al.* (1997) could not

be detected in the amino acid sequences of the wSSII cDNA clones of the present invention, however peptide 1 of Takaoka *et al.* (1997) can be found in the amino acid sequences of SSI from maize, rice, wheat and potato (data not shown).

Denyer *et al.* (1995) demonstrated that the Sgp-1 proteins possess starch synthase activity and, as a consequence, the wSSIIB, wSSIA and wSSIID cDNA clones encode starch synthase enzymes that are differentially expressed in a developmentally-regulated manner in both the soluble and granule-bound fractions of the endosperm (Figure 1). Based on the nomenclature suggested by Harn *et al.* (1998), it is appropriate to describe the Sgp-1 proteins as "starch synthases" rather than "granule-bound starch synthases".

EXAMPLE 10

Analysis of wheat starch synthase II mRNA expression

The mRNA for wheat starch synthase II could be detected in leaves, pre-anthesis florets and endosperm of wheat when total RNAs isolated from these tissue were probed with a PCR probe, wSSIIp2, corresponding to nucleotide positions 1435 to 1835 bp of wSSIIB-cDNA (SEQ ID NO:<400>1; Figure 4). Unlike wSSI, which could not be detected in wheat leaves derived from plants grown under the same conditions, wSSII genes are highly-expressed in the leaves (Figure 4, lane 1), and expressed at an intermediate level in pre-anthesis florets (Figure 4, lane 2), and at much lower levels in developing wheat endosperm cells (Figure 4, lanes 3-11). In contrast, the maize SSIIa is expressed predominantly in the endosperm, whilst the maize SSIIb is detected mainly in the leaf, albeit at low levels (Harn *et al.*, 1998).

The wSSII mRNA was detectable in the endosperm 6 days after anthesis and mRNA levels increase between 8 and 18 days post-anthesis, after which time levels of mRNA decline.

Southern blotting experiments in wheat demonstrated that the wSSIIp2 probe used detected only a single copy of the SSII gene in each genome (data not shown). Thus, it is unlikely that this probe cross-hybridised with mRNAs encoded by genes other than wSSII.

EXAMPLE 11

Chromosomal localization of the wheat wSSII genes.

I. Amplification of specific cDNA regions of wheat starch synthase II using PCR

Two PCR products, wSSIIp2 and wSSIIp3 were amplified from the cDNA clone wSSIIB and used for the northern hybridisation and Southern hybridisation, respectively.

The primers ssIIa (5' TGTTGAGGTTCCATGGCACGTTC 3': SEQ ID NO: <400>25) and ssIIb (5' AGTCGTTCTGCCGTATGATGTCG 3': SEQ ID NO: <400>26) were used to amplify the cDNA fragment wSSIIp2 (i.e. nucleotide positions 1435 to 1835 of SEQ ID NO:<400>1).

The primers ssIIc (5' CCAAGTACCAGTGGTGAACGC 3': SEQ ID NO: <400>27) and ssIId (5' CGGTGGGATCCAACGGCCC 3': SEQ ID NO: <400>28) were used to amplify the cDNA fragment wSSIIp3 (i.e. nucleotide positions 2556 to 2921 of SEQ ID NO:<400>1).

The amplification reactions were performed using a FTS-1 thermal sequencer (Corbett, Australia) for 1 cycle of 95°C for 2 minutes; 35 cycles of 95°C for 30 seconds, 60°C for 1 minutes, 72°C for 2 minutes and 1 cycle of 25°C for 1 minute.

II. PCR and nucleotide sequence analysis of 3' sequences of wheat SSII genes

Genomic DNA was extracted from wild-type Chinese Spring wheat, and from three nullisomic-tetrasomic lines of chromosome 7 of Chinese Spring wheat, and from

Triticum tauschii (var strangulata, accession number CPI 100799), and used as a template for the amplification and nucleotide sequence analysis of wheat SSII genes.

RFLP analysis of *Bam*HI and *Eco*RI restricted DNA from each wheat or *T. Tauschii* line was carried out using the wSSIIp3 fragment as a probe. Three hybridising bands were obtained which could be assigned to chromosomes 7A, 7B and 7D, respectively (data not shown). This analysis indicates that there is a single copy of the wSSII gene in each genome in hexaploid wheat, consistent with the findings of Yamamori and Endo (1996) who located the SGP-A1, B1 and D1 proteins to the short arm of chromosome 7.

PCR analysis was used to assign each of the cDNA clones to the individual wheat genomes. A single 365 bp PCR fragment was obtained from nullisomic-tetrasomic genomic DNA of Chinese Spring when primers ssIIc and ssIId were used for the PCR amplification (Figure 5, right panel). This PCR product is obtained only from lines bearing the B genome. The fragment was cloned and sequenced and shown to be identical to a 365 bp region of the wSSIIB cDNA. An identical fragment is obtained by PCR amplification of the wSSIIB cDNA clone, but not by amplification of the wSSIIA or wSSIID clones, supporting the conclusion that the wSSIIB cDNA is the product of a gene located on chromosome 7 of the B genome of hexaploid wheat.

Two PCR products were also amplified from nullisomic-tetrasomic genomic DNA of Chinese Spring using the primers ssIIc and ssIIe (Figure 5, left panel). One PCR fragment, approximately 350 bp is only amplified when the A genome is present, and a second 322 bp product is only amplified when the D-genome is present. The 350 and 322 bp PCR products were also cloned and sequenced and shown to be identical to the wSSIIA and wSSIID cDNAs, respectively, supporting the conclusion that the wSSIIA and wSSIID cDNAs are the products of genes located on chromosomes 7A and 7D, respectively.

EXAMPLE 12

Isolation of genomic wSSII clones

Screening of a genomic library from the D-genome donor of wheat, *T. tauschii*, was performed as described in Example 5, using the the PCR-derived DNA fragment wSSIIp2 as a hybridisation probe. Figure 6 shows an example of a plaque lift showing the positive-hybridising clone wSSII-8, a putative *T. tauschii* homologue of the wSSII gene.

Positive-hybridising plaques were digested using the restriction enzyme *Bam*HI, separated on a 1% agarose gel, transferred to nitrocellulose membrane and hybridised to probe wSSIIp4 comprising nucleotides 1 to 367 of the wSSIIA cDNA clone, using the conditions described by Rahman *et al.* (1997). As shown in Figure 7,clone wSSII-8 also hybridises strongly to this probe, confirming its identity as a genomic wSSII gene. Furthermore, in light of the fact that the wSSIIp4 probe comprises the 5′-end of the cDNA clone, it is likely that genomic clone wSSII-8 comprises the promoter region of the wSSII gene.

EXAMPLE 13

Cloning of specific cDNA regions of wheat starch synthase III using RT-PCR

PCR primers were used to amplify sequences of starch synthase III from wheat endosperm cDNA. The design of PCR primers was based on the sequences of starch synthase III from potato and the *du1* starch synthase III gene of maize.

First-strand cDNAs were synthesised from 1 μg of total RNA (derived from endosperm of the cultivar Rosella, 12 days after anthesis) as described by Maniatis *et al.* (1982), and then used as templates to amplify two specific cDNA regions, wSSIIIp1 and wSSIIIp2, of wheat starch synthase III by PCR.

The primers used to obtain the cDNA clone wSSIIIp1 were as follows:
Primer wSS3pa (5' GGAGGTCTTGGTGATGTTGT 3': SEQ ID NO: <400>29); and
Primer wSS3pb (5' CTTGACCAATCATGGCAATG 3': SEQ ID NO: <400>30).

The primers used to obtain the cDNA clone wSSIIIp2 were as follows:
Primer wSS3pc (5' CATTGCCATGATTGGTCAAG 3': SEQ ID NO: <400>31); and
Primer wSS3pd (5' ACCACCTGTCCGTTCCGTTGC 3': SEQ ID NO: <400>32).

The amplified clones wSSIIIp1 and wSSIIIp2 were used as probes to screen the third cDNA library and *T. tauschii* genomic DNA library as described in Example 4.

A further probe designated wSSIIIp3 was used for screening the third cDNA library, as described in Example 4. Probe wSSIIIp3 was amplified by PCR from a cDNA clone produced from the first screening using the following amplification primers:
Primer wSS3pe (5' GCACGGTCTATGAGAACAATGGC 3': SEQ ID NO:<400>33); and
Primer wSS3pf (5' TCTGCATACCACCAATCGCCG 3': SEQ ID NO: <400>34).

The amplification reactions were performed using a FTS-1 or FTS4000 thermal sequencer (Corbett, Australia) for 1 cycle of 95°C for 2 minutes; 35 cycles of 95°C for

30 seconds, 60°C for 1 minutes, 72°C for 2 minutes and 1 cycle of 25°C for 1 minute.

Amplified sequences of the expected length were obtained, cloned and sequenced, and shown to contain DNA sequences highly homologous to the maize and potato SSIII genes. PCR fragments were subsequently used to probe a wheat cDNA library by DNA hybridisation and 8 positive clones were obtained, including one 3 kb cDNA. A region from the 5' end of this cDNA was amplified by PCR and used a probe for a second round of screening the cDNA library, obtaining 8 cDNA clones. Of these, one cDNA was demonstrated to be full length (wSSIII.B3, 5.36 kb insert). The sequence of the 5.36 kb wSSIII.B3 cDNA clone is given in SEQ ID NO:<400>7.

Sequencing of the 8 cDNA clones obtained from the second round screening of the wheat cDNA library revealed that there were at least 2 classes of cDNA encoding SSIII present, possibly being encoded by homeologous genes on different wheat genomes. The sequence of a representative of this second class of cDNA clones, wSSIII.B1, is shown in SEQ ID NO:<400>9. The 3664 bp clone wSSIII.B1 is not full length, spanning only the region from nucleotides 1690 to 5363 of the homeologous clone wSSIII.B3, with an open reading frame between nucleotide positions 1 and 3180.

An open reading frame is found between the ATG translation start codon at position 29 and the stop codon at position 4921 of the cDNA clone wSSIII.B3 (SEQ ID NO:<400>7). The amino acid sequence deduced from this open reading frame is shown in SEQ ID NO:<400>8.

An alignment of the deduced amino acid sequences of SSIII from maize, potato and wheat is shown in Figure 8. There is 56.6% identity between the maize SSIII and wheat wSSIII.B3 sequences at the amino acid level.

The C-terminal domain of starch synthases comprise the catalytic domain, and a characteristic amino acid sequence motif KVGGLGDVVTSLSRAVQDLGHNVEV (SEQ

ID NO:<400> 35) in maize, or alternatively KVGGLGDVVTSLSRAIQDLGHTVEV (SEQ ID NO: <400>36) in wheat, marking the first conserved region in the C-terminal domain.

The amino acid identity between maize dull1 and wSSIII.B3 in the N-terminal region (i.e. amino acids 1 to 600 in Figure 8) is only 32.2%; whilst the amino acid identity in the central region (i.e. amino acids 601 to 1248 in Figure 8) is 68.4%; and in the C-terminal region (i.e. amino acids 1249 to 1631 in Figure 8) is 84.6%. Accordingly, the SSIII starch synthases are much more highly conserved between maize and wheat in the region comprising the catalytic domain of the proteins.

EXAMPLE 14

Isolation of genomic clones for SSIII

Screening of a genomic library from the D-genome donor of wheat, *T. tauschii*, identified a number of clones which hybridised to the wSSIII PCR fragment. Positive plaques in the genomic library were selected as those hybridising with a probe that had been generated by PCR (amplifying between nucleotide positions 3620 to 3966) from the SSIII cDNA as template. The primer sequences used were as follows:
wSS3pa (5' GGAGGTCTTGGTGATGTTGT 3': SEQ ID NO: <400>29); and
wSS3pb (5' CTTGACCAATCATGGCAATG 3' : SEQ ID NO: <400>30).

Hybridisation was carried out in 25% formamide, 6 x SSC, 0.1% SDS at 42 °C for 16 hour, then washed three times with 2 x SSC containing 0.1% SDS at 65 °C, for 1 hour per wash. Figure 9 shows an example of a plaque lift showing positive and negative hybridisations for plaques containing the *T. tauschii* homologue of the wSSIII.B3 gene.

DNA was isolated from positive-hybridising λ clones using methods described by Maniatis *et al.* Briefly, DNA was digested using *Bam*HI or *Bgl*I and sub-cloned in to the vector pJKKmfm. DNA sequencing was performed using the automated ABI system with dye terminators as described by the manufacturers. DNA sequences were

analysed using the GCG suite of programs (Devereaux *et al.,* 1984).

Nucleotide equences of the genomic SSIII clone from *T. tauschii* are provided gerein as 6 contiguous sequences designated fragments 1 to 6 (SEQ ID NOs:<400>11 to <400>16, respectively). Table 3 defines the relative positions of these fragments with respect to the SSIII cDNA and describes the positions of exons. Figure 1 shows this information schematically.

EXAMPLE 15

Analysis of wheat starch synthase III mRNA expression

Figure 10 shows the expression of wSSIII mRNA during endosperm development in two wheat varieties grown under defined environmental conditions. The expression of the gene is seen very early in endosperm development in both cultivars, 4 days after anthesis (Figure 10, panels a and b). Expression in the leaf of the variety Gabo is very weak (Figure 10, panel c, Lane L) whereas strong expression is seen in pre-anthesis florets (Figure 10, panel c, Lane P).

TABLE 3

Summary of the Wheat Starch Synthase III Genomic Sequence

Fragment in genomic DNA clone	Length (bp)	Features in SEQ ID NOs:<400>11 to <400>16	Corresponding region in cDNA sequence
Fragment 1 (SEQ ID NO:<400>11)	728	Translation start codon (nucleotides 287 to 289); Exon 1.1 (nucleotides 260 to385).	Exon 1.1: nucleotides 1 to 126
Fragment 2 (SEQ ID NO:<400>12)	2446	Exon 2.1 (nucleotides 1 to 1938); Exon 2.2 (nucleotides 2197 to 2418).	Exon 2.1: nucleotides 1008 to 2948; Exon 2.2: nucleotides 2949 to 3171
Fragment 3 (SEQ ID NO:<400>13)	1032	Exon 3.1 (nucleotides 310 to 580)	Exon 3.1: nucleotides 3172 to 3440
Fragment 4 (SEQ ID NO:<400>14)	892	Exon 4.1 (nucleotides 678 to 853)	Exon 4.1: nucleotides 3441 to 3616
Fragment 5 (SEQ ID NO:<400>15)	871	Partial Exon 5.1 (nucleotides 1 to 29) Exon 5.2 (nucleotides 293 to 463) Exon 5.3 (nucleotides 589 to 695)	Exon 5.1: nucleotides 3908 to 3937 (partial) Exon 5.2: nucleotides 3938 to 4108 Exon 5.3: nucleotides 4109 to 4215
Fragment 6 (SEQ ID NO:<400>16)	1583	Exon 6.1 (nucleotides 471 to 653); Exon 6.2 (nucleotides 770 to 902); Exon 6.3 (nucleotides 999 to 1110); Exon 6.4 (nucleotides 1201 to 1328); Partial Exon 6.5 (nucleotides 1408 to 1583); Translation stop codon (nucleotides 1536 to 1538)	Exon 6.1: nucleotides 4238 to 4420 Exon 6.2: nucleotides 4421 to 4552 Exon 6.3: nucleotides 4553 to 4664 Exon 6.4: nucleotides 4665 to 4793 Exon 6.5: nucleotides 4794 to 4966 (partial)

EXAMPLE 16

Amino acid sequence comparisons between wheat SSII and SSIII polypeptides

Amino acid sequence comparisons between wheat BSSS, SSI, SSII and SSIII polypeptides, as indicated in Figure 13, reveals eight highly-conserved domains. The amino acid sequences of these domains are represented in the wheat SSIII amino acid sequence by the following sequence motifs:

(a) Region1: KVGGLGDVVTS;
(b) Region 2: GHTVEVILPKY;
(c) Region 3: HDWSSAPVAWLYKEHY;
(d) Region 4: GILNGIDPDIWDPYTD;
(e) Region 5: DVPIVGIITRLTAQKG;
(f) Region 5a: NGQVVLLGSA;
(g) Region 6: AGSDFIIVPSIFEPCGLTQLVAMRYGS; and
(h) Region 7: TGGLVDTV.

As shown in Table 4 below, there is at least about 25% amino acid sequence identity, preferably at least about 30% amino acid sequence identity, more preferably at least about 35% amino acid sequence identity, more preferably at least about 40% amino acid sequence identity, more preferably at least about 45% amino acid sequence identity, more preferably at least about 50% amino acid sequence identity, more preferably at least about 55% amino acid sequence identity, more preferably at least about 60% amino acid sequence identity, more preferably at least about 65% amino acid sequence identity, more preferably at least about 70% amino acid sequence identity, more preferably at least about 75% amino acid sequence identity, more preferably at least about 80% amino acid sequence identity, more preferably at least about 85% amino acid sequence identity, more preferably at least about 90% amino acid sequence identity and even more preferably at least about 95% amino acid sequence identity between the

amino acid sequences of plant starch synthase enzymes, in particular wheat starch synthases.

TABLE 4

Identities between conserved motifs of plant starch synthases

Conserved Region	Number of conserved residues between wheat starch synthases	Number of conserved residues between wheat SSII and SSIII polypeptides
Region1: KVGGLGDVVTS	6/11 residues	6/11 residues
Region 2: GHTVEVILPKY	6/11 residues	6/11 residues
Region 3: HDWSSAPVAWLYKEHY	4/16 residues	5/16 residues
Region 4: GILNGIDPDIWDPYTD	7/16 residues	8/16 residues
Region 5: DVPIVGIITRLTAQKG	8/16 residues	10/16 residues
Region 5a: NGQVVLLGSA	4/10 residues	4/10 residues
Region 6: AGSDFIIVPSIFEPCGLT QLVAMRYGS	15/27 residues	17/27 residues
Region 7: TGGLVDTV	5/9 residues	5/9 residues

The most conserved regions of the wheat SSII and SSIII polypeptides are a region of 6 or 7 identical amino acids in Region 1 (Table 4; Figure 13) and a region of 8 or

9 identical amino acids in Region 6 (Table 4; Figure 13). The lowest regions of identity are found in regions 3 and 5a.

EXAMPLE 17

Discussion

Early work on the Sgp-1 starch synthase proteins (Denyer *et al.*, 1995; Rahman *et al.*, 1995) was based on the localisation of these proteins in the wheat starch granule, and no definitive conclusion concerning their presence or absence in soluble extracts of the wheat endosperm was presented. We have now demonstrated that a monoclonal antibody against the Sgp-1 proteins cross reacts strongly with those starch synthase proteins having apparent molecular weights of 100-105 kDa in soluble extracts, however, the appearance of these proteins in soluble extracts is dependant on the developmental stage of the endosperm material. Whilst the proteins can be detected in the soluble phase in early to mid endosperm development, little or no soluble protein remains in late endosperm development (Figure 1). This observation accounts for the failure of Rahman *et al.* (1995) to detect the protein in soluble extracts in a previous report.

Based upon the localisation of the Sgp-1 starch synthase proteins in the wheat endosperm, the following nomenclature is suggested for wheat starch synthase enzymes: wGBSS for the 60 kDa granule bound starch synthase (Wx); wSSI for the 75 kDa starch synthase I (Sgp-3); wSSII for the 100 - 105 kDa proteins (Sgp-1); and wSSIII for a soluble high molecular starch synthase.

The present invention provides cDNA clones encoding the wSSII and wSSIII polypeptides and the corresponding genomic clones.

The wSSIII cDNA clone described herein is clearly related to the maize and potato SSIII polypeptides.

Comparison of the amino acid sequences of all available starch synthases with the deduced amino acid sequences of the three wSSII cDNA clones of the present invention (i.e. wSSIIB, wSSIIA and wSSIID) was conducted using PILEUP analysis (Devereaux *et al.*, 1984) and data are presented as a dendrogram (Figure 11). The sequence of the glycogen synthase of *E. coli* was also included. Based upon their amino acid similarities, four classes of plant starch synthases can be defined: GBSS, SSI, SSII and SSIII.

Based upon sequence identities and the function of the Sgp-1 proteins in wheat, the wSSIIB, wSSIIA and wSSID cDNA clones are members of the starch synthase II (SSII) group and are more similar in sequence to maize SSIIa than maize SSIIb. Table 5 shows that levels of identity at the amino acid level between the wSSII sequences, as determined using the BESTFIT programme in GCG (Devereaux *et al.*, 1984), and other class II starch synthases range from 70% identity with potato SSII to 85% identity with maize SSIIa. Both wSSIIB and wSSIID showed significantly higher homology to maize SSIIa than wSSIIA.

TABLE 5

	wSSII-A	**wSSII-B**	**wSSII-D**
wSSI-A	100%		
wSSII-B	95.9%	100%	
wSSII-D	96.3%	96.7%	100%
maize SSIIa	76.1%	85.2%	84.7%
maize SSIIb	76.3%	76.7%	75.9%
pea SSII	72.0%	72.2%	71.8%
potato SSII	70.9%	71.1%	70.3%

Whilst the evidence is compelling that the wSSIIA, wSSIIB and wSSIID cDNAs encode the Sgp-A1, Sgp-B1 and Sgp-D1 proteins of the wheat starch granule, molecular weights calculated from the deduced amino acid sequences of the clones

are considerably lower than estimates obtained from SDS-PAGE. The molecular weight of the precursor wSSIIA protein is 87,229 Da, and the mature protein 81,164 Da, yet the estimated molecular weight in our experience is 105 kDa. The assignment of the wSSIIA cDNA to the A-genome of wheat is demonstrated in Figure 5, and the assignement of the 105 kDa protein to the A-genome in Denyer *et al.* (1995) and Yamamori and Endo (1996). Similarly, the molecular weight of the wSSIIB protein is 86,790 Da and the mature protein 80,759 Da, yet the molecular weight of the Sgp-B1 protein is estimated to be 100 kDa. No comparison can be made of the wSSIID sequences as a full length cDNA clone was not obtained. The wSSIIA and wSSIIB amino acid sequences differ by just a single amino acid residue, yet there is an apparent difference of 5 kDa in molecular weight when estimated by SDS-PAGE. Several possibilities can be advanced to account for this apparent discrepancy in molecular weights. Firstly, the wSSII proteins may not migrate in SDS-PAGE in accordance with their molecular weight because they retain some conformation under the denaturing conditions used. Secondly, the proteins may be glycosylated. It is also possible that the proteins may be non-covalently linked to starch through a high affinity starch binding site which survives denaturation and SDS-PAGE. Differences between the apparent molecular weights and those calculated from the deduced amino acid sequences will have to be defined in establishing the relationship between the wSSII proteins and proteins encoded by the analogous SSII genes of other species.

The catalytic domain of the starch synthases is found at the C-terminal end of the protein (Gao *et al.*, 1998; Harn *et al.*, 1998). Harn *et al.* (1998) identified 7 conserved regions among SSIIa, SSIIb, SSI and GBSS sequences.

We have identified include an additional conserved region (designated region 5a in Figure 3 and Figure 12) comprising the amino acid sequence motif DVQLVMLGTG. Comparison of the wSSII sequences of the present invention with differing isoforms of starch synthases (GBSS, SS1, SSII and SSIII) identified a total of 8 regions of

the deduced amino acid sequences which were conserved amongst starch synthases from each class. Figure 12 shows an alignment of plant starch synthase sequences, in which the position of the first homologous region is used as the basis of the alignment. This first homologous region contains the consensus motif KXGG which is believed to be present in the ADPglucose binding site of starch and glycogen synthases (Furukawa *et al.*, 1990).

The conservation of eight peptide regions among the 4 classes of starch synthases is striking, in terms of their sequence homologies and their alignment. The major differences in structure between the classes of genes are found in the length of the N-terminal region between the transit peptide and the first conserved region. At one extreme, the GBSS genes have a very short N-terminal arm, whereas the *du1* starch synthase contains a very long N-terminal extension containing several distinct regions (Figure 12). The wSSII genes contain an N-terminal extension which is longer than either GBSS,SSI, or SSIIb, and slightly longer than the maize SSIIa gene (Figure 12). Analysis of the wheat SSII genes shows that there is a motif, PVNGENK, which is repeated. The area surrounding the repeated PVNGENK motif is not homologous to maize SSIIa and the insertion of this region is responsible for the difference in length between the wheat SSII and maize SSIIa genes. In pea and potato SSII polypeptides, a PPP motif (Figure 3; residues 251-253 and 287-289 respectively) has been suggested to mark the end of the N-terminal region and to facilitate the flexibility of an “N-terminal arm”. This motif is not found in either the maize or wheat SSII sequences.

The generation of a wheat line combining null alleles at each of the three wSSII loci, wSSIIA, wSSIIB and wSSIID, has been reported recently by Yamamori (1998). In this triple null line, the large starch granules were reported to be mostly deformed and a novel starch with high blue value was observed when stained with iodine, indicating that wSSII is a key enzyme for the synthesis of starch in wheat. Further analysis of the starch derived from this triple null mutant is in progress.

Mutations in starch synthases are known in three other species. In pea, mutation in SSII gives rise to starch with altered granule morphology and an amylopectin which yields an oligosaccharide distribution with reduced chain length on debranching, compared to the wild type (Craig *et al.*, 1998). A similar mutation in a gene designated SSII is known in *Chlamydomonas* (the *sta*-3 mutation) and similar effects on granule morphology and amylopectin structure are observed (Fontaine *et al.*, 1993). In maize, two mutations affecting starch synthases are known. First, the *dull1*mutation has been shown to be caused by a lesion within the *du*1 SSIII-type starch synthase gene (Gao *et al.*, 1998). A second mutation, the *sugary*-2 mutation yields a starch with reduced amylopectin chain lengths on debranching (this mutation co-segregates with the SSIIa locus (Harn *et al.*, 1998) although direct evidence that the *sugary*-2 mutation is caused by a lesion in the SSIIa gene is lacking). In the SSII mutants of each of these species, amylose biosynthesis capacity is retained, suggesting different roles in amylose and amylopectin synthesis for the GBSS and SSII genes. Given the conservation in overall organisation of the GBSS and SSII genes (see Figures 11 and 12), when an alignment is made based on the KTGGL motif of the first conserved region, this focuses attention on the role(s) of the N-terminal region in defining substrate specificity and the localisation of the proteins as the N-terminal region is the major area of divergence between the 4 classes of starch synthases. However, it is premature to exclude the influence of more subtle mutations in central and C-terminal regions of the gene.

The cloning of the wSSII and wSSIII cDNAs and genomic clones described herein provides useful tools for the further study of the roles of the starch synthases in wheat. Firstly, they provide a source of markers which can be used to recover and combine null or divergent alleles. Secondly, genetic manipulation of wheat by gene suppression or over-expression can be carried out, and the genes may be used for overexpression in other species. The promoter regions of these genes are also useful in regulating the expression of starch synthase genes and other

heterologous genes in the developing wheat endosperm and in pre-anthesis florets of wheat.

REFERENCES

1. Ausubel, F. M., Brent, R., Kingston, RE, Moore, D.D., Seidman, J.G., Smith, J.A., and Struhl, K. (1987). *In:* Current Protocols in Molecular Biology. Wiley Interscience (ISBN 047150338).

2. Abel GJW, Springer F, Willmitzer L, Kossmann J (1996) Cloning and functional analysis of a cDNA encoding a novel 139 kDa starch synthase from potato (*Solanum tuberosum* L.). Plant J **10**: 981-991.

3. Ainsworth C, Clark J, Balsdon J (1993) Expression, organisation and structure of the genes encoding the waxy protein (granule-bound starch synthase) in wheat. Plant Mol Biol **22:** 67-82.

4. Baba T, Nishihara M, Mizuno K, Kawasaki T, Shimada H, Kobayabashi E, Ohnishi S, Tanaka K, Arai Y (1993) Identification, cDNA cloning, and Gene Expression of Soluble Starch Synthase in Rice (*Oryza sativa* L.) Immature Seeds. Plant Physiol **103**: 565-573.

5. Craig J, Lloyd JR, Tomlinson K, Barber L, Edwards A, Wang TL, Martin C, Hedley CL, Smith AM (1998) Mutations in the gene encoding starch synthase II profoundly alter amylopectin structure in pea embryos. Plant Cell **10**: 413-426.

6. Denyer K, Hylton CM, Jenner CF, Smith AM (1995) Identification of multiple isoforms of soluble and granule-bound starch synthase in developing wheat endosperm. Planta **196**: 256-265.

7. Devereaux, J, Haeberli P, Smithies O (1984) A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res **12**: 387-395.

8. Dry I, Smith A, Edwards A, Bhattacharyya M, Dunn P, Martin C (1992) Characterisation of cDNAs encoding two isoforms of granule-bound starch synthase which show differential expression in developing storage organ of pea and potato. Plant J **2**: 193-202.

9. Edwards A, Marshall J, Sidebottom C, Visser RGF, Smith AM, Martin C (1995) Biochemical and molecular characterization of a novel starch synthase from potato tubers. Plant J **8**: 283-294.

10. Fontaine T, D'Hulst C, Maddelein M-L, Routier F, Pépin TM, Decq A, Wieruszeski J-M, Delrue B, Van den Koornhuyse N, Bossu J-P, Fournet B, Ball S (1993) Toward an understanding of the biogenesis of the starch granule. Evidence that *Chlamydomonas* soluble starch synthase II controls the synthesis of intermediate size glucans of amylopectin. J Biol Chem **22**: 16223-16230.

11. Furukawa K, Tagaya M, Inouye M, Preiss J, Fukui T (1990) Identification of lysine 15 at the active site in *Escherichia coli* glycogen synthase. J Biol Chem **265**: 2086-2090.

12. Gao M, Wanat J, Stinard PS, James MG, Myers AM (1998) Characterization of dull1, a maize gene coding for a novel starch synthase. Plant Cell **10**: 399-412.

13. Harn C, Knight M, Ramakrishnan A, Guan H, Keeling PL, Wasserman BP (1998) Isolation and characterization of the *zSSIIa* and *zSSIIb* starch synthase cDNA clones from maize endosperm. Plant Mol Biol **37**: 639-649.

14. Kloesgen RB, Gierl A, Schwarz-Sommer ZS, Saedler H (1986) Molecular analysis of the waxy locus of Zea mays. Mol Gen Genet **203**: 237-244.

15. Knight ME, Harn C, Lilley CER, Guan H, Singletary G W, MuForster C, Wasserman BP, Keeling PL (1998) Molecular cloning of starch synthase I from maize (W64) endosperm and expression in *Escherichia*. Plant J **14**: 613-622.

16. Kumar A, Larsen CE, Preiss J (1986) Biosynthesis of bacterial glycogen: Primary structure of *Escherichia coli* ADP-glucose:alpha-1,4-glucan, 4-glucosyltransferase as deduced from the nucleotide sequence of the *glg*A gene. J Biol Chem **261**: 16256-16259.

17. Li Z, Rahman S, Kosar-Hashemi B, Mouille G, Appels R Morell, MK (1999)

Cloning and characterisation of a gene encoding wheat starch synthase I. Theor Appl Genet: *In press.*

18. Mouille G, Maddelein M-L, Libessart N, Talaga P, Decq A, Delrue B Ball, S (1996). Preamylopectin processing: A mandatory step for starch biosynthesis in plants. Plant Cell **8**: 1353-1366.

19. Nakamura T, Yamamori M, Hirano H, Hidaka S, Nagamine T (1995) Production of waxy (amylose-free) wheats. Mol Gen Genet **248**: 253-259.

20. Okagaki, RJ (1992) Nucleotide sequence of a long cDNA from the rice waxy gene. Plant Mol Biol **19**: 513-516.

21. Ozbun, J.L., Hawker, J.S. and Preiss, J. (1971) Adensine diphosphoglucose-starch glucosyltransferases from developing kernels of waxy maize. Plant Physiology **48**: 765-769

22. Ozbun, J.L., Hawker, J.S., Greenberg, E., Lammel, C., Preiss, J. and Lee, E.Y.C.(1973) Starch synthetase, phosphorylase, ADPglucose pyrophosphorylase, and UDPglucose pyrophosphorylase in developing maize kernels. Plant Physiology **51**: 1-5.

23. Pollock, C. and Preiss, J.(1980) The citrate-stimulated starch synthase of starchy maize kernels: purification and properties. Arch Biochem Biophys **204**: 578-588.

24. Rahman S, Abrahams S, Abbott D, Mukai Y, Samuel M, Morell M, Appels R (1997) A complex arrangement of genes at a starch branching enzyme I locus in D-genome donor of wheat. Genome **40**: 465-474.

25. Rahman S, Kosar-Hashemi B, Samuel M, Hill A, Abbott DC, Skerritt JH, Preiss J, Appels R, Morell M (1995) The major proteins of wheat endosperm starch granules. Aust J Plant Physiol **22**: 793-803.

26. Rahman S, Li Z, Abrahams S, Abbott D, Appels R, Morell M (1998) Characterisation of a gene encoding wheat endosperm starch branching enzyme-I. Theor Appl Genet **98**: *In press*.

27. Sears ER, Miller TG (1985) The history of Chinese spring wheat. Cereal Res Comm **13**: 261-263.

28. Takaoka M, Watanabe S, Sassa H, Yamamori M, Nakamura T, Sasakuma T, Hirano H (1997) Structural characterisation of high molecular weight starch granule-bound proteins in wheat (*Triticum aestivum* L). J Agric Food Chem **45**: 2929-2934.

29. van der Leij FR, Visser RGF, Ponstein AS, Jacobsen E, Feenstra WJ (1991) Sequence of the structural gene for granule bound starch synthase of potato (*Solanum tuberosum* L.) and evidence for a single point deletion in the *amf* allele. Mol Gen Genet **228**: 240-248.

30. Yamamori M, Endo TR (1996) Variation of starch granule proteins and chromosome mapping of their coding genes in common wheat. Theor Appl Genet **93**: 275-281.

31. Yamamori M (1998) Selection of a wheat lacking a putative enzyme for starch synthesis, SGP-1 *Proc 9th Int Wheat Gen Symp* **4**, 300-302.

SEQUENCE LISTING

```
<110> COMMONWEALTH SCIENTIFIC AND INDUSTRIAL ORGANISATION

<120> NOVEL GENES ENCODING WHEAT STARCH SYNTHASES AND USES THEREFOR

<130> p:\oper\mro\pi-wss.prv

<140>
<141>

<160> 36

<170> PatentIn Ver. 2.0

<210> 1
<211> 2939
<212> DNA
<213> Triticum aestivum

<220>
<221> CDS
<222> (176)..(2572)

<400> 1
atttcctcgg cctgaccccg tgcgtttacc ccacacagag cacactccag tccagtccag 60

cccactgccg cgctactccc cactcccact gccaccacct ccgcctgcgc cgcgctctgg 120

gcggaccaac ccgcgcatcg tatcacgatc acccaccccg atcccggccg ccgcc atg  178
                                                            Met
                                                             1

tcg tcg gcg gtc gcg tcc gcc gcg tcc ttc ctc gcg ctc gcg tcc gcc    226
Ser Ser Ala Val Ala Ser Ala Ala Ser Phe Leu Ala Leu Ala Ser Ala
              5                  10                  15

tcc ccc ggg aga tca cgg agg agg acg agg gtg agc gcg tcg cca ccc    274
Ser Pro Gly Arg Ser Arg Arg Arg Thr Arg Val Ser Ala Ser Pro Pro
         20                  25                  30

cac acc ggg gct ggc agg ttg cac tgg ccg ccg tcg ccg ccg cag cgc    322
His Thr Gly Ala Gly Arg Leu His Trp Pro Pro Ser Pro Pro Gln Arg
     35                  40                  45

acg gct cgc gac gga gcg gtg gcc gcg cgc gcc gcc ggg aag aag gac    370
Thr Ala Arg Asp Gly Ala Val Ala Ala Arg Ala Ala Gly Lys Lys Asp
 50                  55                  60                  65

gcg ggg atc gac gac gcc gcg ccc gcg agg cag ccc cgc gca ctc cgc    418
Ala Gly Ile Asp Asp Ala Ala Pro Ala Arg Gln Pro Arg Ala Leu Arg
                 70                  75                  80

ggt ggc gcc gcc acc aag gtt gcg gag cgg agg gat ccc gtc aag acg    466
Gly Gly Ala Ala Thr Lys Val Ala Glu Arg Arg Asp Pro Val Lys Thr
             85                  90                  95

ctc gat cgc gac gcc gcg gaa ggt ggc gcg ccg tcc ccg ccg gca ccg    514
Leu Asp Arg Asp Ala Ala Glu Gly Gly Ala Pro Ser Pro Pro Ala Pro
        100                 105                 110

agg cag gag gac gcc cgt ctg ccg agc atg aac ggc atg ccg gtg aac    562
Arg Gln Glu Asp Ala Arg Leu Pro Ser Met Asn Gly Met Pro Val Asn
    115                 120                 125
```

```
ggt gaa aac aaa tct acc ggc ggc ggc ggc gcg act aaa gac agc ggg    610
Gly Glu Asn Lys Ser Thr Gly Gly Gly Gly Ala Thr Lys Asp Ser Gly
130                 135                 140                 145

ctg ccc gca ccc gca cgc gcg ccc cag ccg tcg agc cag aac aga gta    658
Leu Pro Ala Pro Ala Arg Ala Pro Gln Pro Ser Ser Gln Asn Arg Val
                150                 155                 160

ccg gtg aat ggt gaa aac aaa gct aac gtc gcc tcg ccg ccg acg agc    706
Pro Val Asn Gly Glu Asn Lys Ala Asn Val Ala Ser Pro Pro Thr Ser
            165                 170                 175

ata gcc gag gtc gcg gct ccg gat ccc gca gct acc att tcc atc agt    754
Ile Ala Glu Val Ala Ala Pro Asp Pro Ala Ala Thr Ile Ser Ile Ser
        180                 185                 190

gac aag gcg cca gag tcc gtt gtc cca gcc gag aag gcg ccg ccg tcg    802
Asp Lys Ala Pro Glu Ser Val Val Pro Ala Glu Lys Ala Pro Pro Ser
    195                 200                 205

tcc ggc tca aat ttc gtg ccc tcg gct tct gct ccc ggg tct gac act    850
Ser Gly Ser Asn Phe Val Pro Ser Ala Ser Ala Pro Gly Ser Asp Thr
210                 215                 220                 225

gtc agc gac gtg gaa ctt gaa ctg aag aag ggt gcg gtc att gtc aaa    898
Val Ser Asp Val Glu Leu Glu Leu Lys Lys Gly Ala Val Ile Val Lys
                230                 235                 240

gaa gct cca aac cca aag gct ctt tcg ccg ccc gca gca ccc gct gta    946
Glu Ala Pro Asn Pro Lys Ala Leu Ser Pro Pro Ala Ala Pro Ala Val
            245                 250                 255

caa caa gac ctt tgg gac ttc aag aaa tac att ggt ttc gag gag ccc    994
Gln Gln Asp Leu Trp Asp Phe Lys Lys Tyr Ile Gly Phe Glu Glu Pro
        260                 265                 270

gtg gag gcc aag gat gat ggc cgg gct gtt gca gat gat gcg ggc tcc    1042
Val Glu Ala Lys Asp Asp Gly Arg Ala Val Ala Asp Asp Ala Gly Ser
    275                 280                 285

ttc gaa cac cac cag aat cac gat tcc ggg cct ttg gca ggg gag aac    1090
Phe Glu His His Gln Asn His Asp Ser Gly Pro Leu Ala Gly Glu Asn
290                 295                 300                 305

gtc atg aac gtg gtc gtc gtg gct gct gaa tgt tct ccc tgg tgc aaa    1138
Val Met Asn Val Val Val Val Ala Ala Glu Cys Ser Pro Trp Cys Lys
                310                 315                 320

aca ggt ggt ctt gga gat gtt gcc ggt gct ttg ccc aag gct ttg gcg    1186
Thr Gly Gly Leu Gly Asp Val Ala Gly Ala Leu Pro Lys Ala Leu Ala
            325                 330                 335

aag aga gga cat cgt gtt atg gtt gtg gta cca agg tat ggg gac tat    1234
Lys Arg Gly His Arg Val Met Val Val Val Pro Arg Tyr Gly Asp Tyr
        340                 345                 350

gag gaa gcc tac gat gtc gga gtc cga aaa tac tac aag gct gct gga    1282
Glu Glu Ala Tyr Asp Val Gly Val Arg Lys Tyr Tyr Lys Ala Ala Gly
    355                 360                 365

cag gat atg gaa gtg aat tat ttc cat gct tat atc gat gga gtt gat    1330
Gln Asp Met Glu Val Asn Tyr Phe His Ala Tyr Ile Asp Gly Val Asp
370                 375                 380                 385
```

ttt gtg ttc att gac gct cct ctc ttc cga cac cgc cag gaa gac att 1378
Phe Val Phe Ile Asp Ala Pro Leu Phe Arg His Arg Gln Glu Asp Ile
390 395 400

tat ggg ggc agc aga cag gaa att atg aag cgc atg att ttg ttc tgc 1426
Tyr Gly Gly Ser Arg Gln Glu Ile Met Lys Arg Met Ile Leu Phe Cys
405 410 415

aag gcc gct gtc gag gtt cca tgg cac gtt cca tgc ggc ggt gtc cct 1474
Lys Ala Ala Val Glu Val Pro Trp His Val Pro Cys Gly Gly Val Pro
420 425 430

tat ggg gat gga aat ctg gtg ttt att gca aat gat tgg cac acg gca 1522
Tyr Gly Asp Gly Asn Leu Val Phe Ile Ala Asn Asp Trp His Thr Ala
435 440 445

ctc ctg cct gtc tat ctg aaa gca tat tac agg gac cat ggt ttg atg 1570
Leu Leu Pro Val Tyr Leu Lys Ala Tyr Tyr Arg Asp His Gly Leu Met
450 455 460 465

cag tac act cgg tcc att atg gtg ata cat aac atc gct cac cag ggc 1618
Gln Tyr Thr Arg Ser Ile Met Val Ile His Asn Ile Ala His Gln Gly
470 475 480

cgt ggc cca gta gat gag ttc ccg ttc acc gag ttg cct gag cac tac 1666
Arg Gly Pro Val Asp Glu Phe Pro Phe Thr Glu Leu Pro Glu His Tyr
485 490 495

ctg gaa cac ttc aga ctg tac gac ccc gtg ggt ggt gaa cac gcc aac 1714
Leu Glu His Phe Arg Leu Tyr Asp Pro Val Gly Gly Glu His Ala Asn
500 505 510

tac ttc gcc gcc ggc ctg aag atg gcg gac cag gtt gtc gtc gtg agc 1762
Tyr Phe Ala Ala Gly Leu Lys Met Ala Asp Gln Val Val Val Val Ser
515 520 525

ccg ggg tac ctg tgg gag ctg aag acg gtg gag ggc ggc tgg ggg ctt 1810
Pro Gly Tyr Leu Trp Glu Leu Lys Thr Val Glu Gly Gly Trp Gly Leu
530 535 540 545

cac gac atc ata cgg cag aac gac tgg aag acc cgc ggc atc gtg aac 1858
His Asp Ile Ile Arg Gln Asn Asp Trp Lys Thr Arg Gly Ile Val Asn
550 555 560

ggc atc gac aac atg gag tgg aac ccc gag gtg gac gtc cac ctc aag 1906
Gly Ile Asp Asn Met Glu Trp Asn Pro Glu Val Asp Val His Leu Lys
565 570 575

tcg gac ggc tac acc aac ttc tcc ctg ggg acg ctg gac tcc ggc aag 1954
Ser Asp Gly Tyr Thr Asn Phe Ser Leu Gly Thr Leu Asp Ser Gly Lys
580 585 590

cgg cag tgc aag gag gcc ctg cag cgg gag ctg ggc ctg cag gtc cgc 2002
Arg Gln Cys Lys Glu Ala Leu Gln Arg Glu Leu Gly Leu Gln Val Arg
595 600 605

ggc gac gtg ccg ctg ctc ggc ttc atc ggg cgc ctg gac ggg cag aag 2050
Gly Asp Val Pro Leu Leu Gly Phe Ile Gly Arg Leu Asp Gly Gln Lys
610 615 620 625

ggc gtg gag atc atc gcg gac gcg atg ccc tgg atc gtg agc cag gac 2098
Gly Val Glu Ile Ile Ala Asp Ala Met Pro Trp Ile Val Ser Gln Asp
630 635 640

gtg cag ctg gtc atg ctg ggc acc ggg cgc cac gac ctg gag ggc atg 2146

```
Val Gln Leu Val Met Leu Gly Thr Gly Arg His Asp Leu Glu Gly Met
            645                 650                 655

ctg cgg cac ttc gag cgg gag cac cac gac aag gtg cgc ggg tgg gtg     2194
Leu Arg His Phe Glu Arg Glu His His Asp Lys Val Arg Gly Trp Val
        660                 665                 670

ggg ttc tcc gtg cgg ctg gcg cac cgg atc acg gcc ggc gcc gac gcg     2242
Gly Phe Ser Val Arg Leu Ala His Arg Ile Thr Ala Gly Ala Asp Ala
    675                 680                 685

ctc ctc atg ccc tcc cgg ttc gag ccg tgc gga ctg aac cag ctc tac     2290
Leu Leu Met Pro Ser Arg Phe Glu Pro Cys Gly Leu Asn Gln Leu Tyr
690                 695                 700                 705

gcc atg gcc tac ggc acc gtc ccc gtc gtg cat gcc gtc ggt ggc ctg     2338
Ala Met Ala Tyr Gly Thr Val Pro Val Val His Ala Val Gly Gly Leu
                710                 715                 720

agg gac acc gtg ccg ccg ttc gac ccc ttc aac cac tcc ggg ctc ggg     2386
Arg Asp Thr Val Pro Pro Phe Asp Pro Phe Asn His Ser Gly Leu Gly
            725                 730                 735

tgg acg ttc gac cgc gca gag gcg cag aag ctg atc gag gcg ctc ggg     2434
Trp Thr Phe Asp Arg Ala Glu Ala Gln Lys Leu Ile Glu Ala Leu Gly
        740                 745                 750

cac tgc ctc cgc acc tac cgg gac tac aag gag agc tgg agg ggg ctc     2482
His Cys Leu Arg Thr Tyr Arg Asp Tyr Lys Glu Ser Trp Arg Gly Leu
    755                 760                 765

cag gag cgc ggc atg tcg cag gac ttc agc tgg gag cat gcc gcc aag     2530
Gln Glu Arg Gly Met Ser Gln Asp Phe Ser Trp Glu His Ala Ala Lys
770                 775                 780                 785

ctc tac gag gac gtc ctc gtc aag gcc aag tac cag tgg tga             2572
Leu Tyr Glu Asp Val Leu Val Lys Ala Lys Tyr Gln Trp
                790                 795

acgctagctg ctagccggtc cagccccgca tgcgtgcatg acaggatgga attgcgcatt 2632

gcgcacgcag gaaggtgcca tggagcgccg gcatccgcga agtacagtga catgaggtgt 2692

gtgtggttga gacgctgatt ccgatctggt ccgtagcaga gtagagcgga ggtagggaag 2752

cgctccttgt tacaggtata tgggaatgtt gttaacttgg tattgtaatt tgttatgttg 2812

tgtgcattat tacagagggc aacgatctgc gccggcgcac cggcccaact gttgggccgg 2872

tcgcacagca gccgttggat ccgaccgcct gggccgttgg atcccaccga aaaaaaaaaa 2932

aaaaaaa                                                           2939


<210> 2
<211> 798
<212> PRT
<213> Triticum aestivum

<400> 2
Met Ser Ser Ala Val Ala Ser Ala Ala Ser Phe Leu Ala Leu Ala Ser
 1               5                  10                  15

Ala Ser Pro Gly Arg Ser Arg Arg Arg Thr Arg Val Ser Ala Ser Pro
            20                  25                  30
```

Pro His Thr Gly Ala Gly Arg Leu His Trp Pro Pro Ser Pro Pro Gln
35 40 45

Arg Thr Ala Arg Asp Gly Ala Val Ala Ala Arg Ala Ala Gly Lys Lys
50 55 60

Asp Ala Gly Ile Asp Asp Ala Ala Pro Ala Arg Gln Pro Arg Ala Leu
65 70 75 80

Arg Gly Gly Ala Ala Thr Lys Val Ala Glu Arg Arg Asp Pro Val Lys
85 90 95

Thr Leu Asp Arg Asp Ala Ala Glu Gly Gly Ala Pro Ser Pro Pro Ala
100 105 110

Pro Arg Gln Glu Asp Ala Arg Leu Pro Ser Met Asn Gly Met Pro Val
115 120 125

Asn Gly Glu Asn Lys Ser Thr Gly Gly Gly Gly Ala Thr Lys Asp Ser
130 135 140

Gly Leu Pro Ala Pro Ala Arg Ala Pro Gln Pro Ser Ser Gln Asn Arg
145 150 155 160

Val Pro Val Asn Gly Glu Asn Lys Ala Asn Val Ala Ser Pro Pro Thr
165 170 175

Ser Ile Ala Glu Val Ala Ala Pro Asp Pro Ala Ala Thr Ile Ser Ile
180 185 190

Ser Asp Lys Ala Pro Glu Ser Val Val Pro Ala Glu Lys Ala Pro Pro
195 200 205

Ser Ser Gly Ser Asn Phe Val Pro Ser Ala Ser Ala Pro Gly Ser Asp
210 215 220

Thr Val Ser Asp Val Glu Leu Glu Leu Lys Lys Gly Ala Val Ile Val
225 230 235 240

Lys Glu Ala Pro Asn Pro Lys Ala Leu Ser Pro Pro Ala Ala Pro Ala
245 250 255

Val Gln Gln Asp Leu Trp Asp Phe Lys Lys Tyr Ile Gly Phe Glu Glu
260 265 270

Pro Val Glu Ala Lys Asp Asp Gly Arg Ala Val Ala Asp Asp Ala Gly
275 280 285

Ser Phe Glu His His Gln Asn His Asp Ser Gly Pro Leu Ala Gly Glu
290 295 300

Asn Val Met Asn Val Val Val Val Ala Ala Glu Cys Ser Pro Trp Cys
305 310 315 320

Lys Thr Gly Gly Leu Gly Asp Val Ala Gly Ala Leu Pro Lys Ala Leu
325 330 335

Ala Lys Arg Gly His Arg Val Met Val Val Val Pro Arg Tyr Gly Asp
340 345 350

Tyr Glu Glu Ala Tyr Asp Val Gly Val Arg Lys Tyr Tyr Lys Ala Ala
355 360 365

Gly Gln Asp Met Glu Val Asn Tyr Phe His Ala Tyr Ile Asp Gly Val

370 375 380

Asp Phe Val Phe Ile Asp Ala Pro Leu Phe Arg His Arg Gln Glu Asp
385 390 395 400

Ile Tyr Gly Gly Ser Arg Gln Glu Ile Met Lys Arg Met Ile Leu Phe
405 410 415

Cys Lys Ala Ala Val Glu Val Pro Trp His Val Pro Cys Gly Gly Val
420 425 430

Pro Tyr Gly Asp Gly Asn Leu Val Phe Ile Ala Asn Asp Trp His Thr
435 440 445

Ala Leu Leu Pro Val Tyr Leu Lys Ala Tyr Tyr Arg Asp His Gly Leu
450 455 460

Met Gln Tyr Thr Arg Ser Ile Met Val Ile His Asn Ile Ala His Gln
465 470 475 480

Gly Arg Gly Pro Val Asp Glu Phe Pro Phe Thr Glu Leu Pro Glu His
485 490 495

Tyr Leu Glu His Phe Arg Leu Tyr Asp Pro Val Gly Gly Glu His Ala
500 505 510

Asn Tyr Phe Ala Ala Gly Leu Lys Met Ala Asp Gln Val Val Val Val
515 520 525

Ser Pro Gly Tyr Leu Trp Glu Leu Lys Thr Val Glu Gly Gly Trp Gly
530 535 540

Leu His Asp Ile Ile Arg Gln Asn Asp Trp Lys Thr Arg Gly Ile Val
545 550 555 560

Asn Gly Ile Asp Asn Met Glu Trp Asn Pro Glu Val Asp Val His Leu
565 570 575

Lys Ser Asp Gly Tyr Thr Asn Phe Ser Leu Gly Thr Leu Asp Ser Gly
580 585 590

Lys Arg Gln Cys Lys Glu Ala Leu Gln Arg Glu Leu Gly Leu Gln Val
595 600 605

Arg Gly Asp Val Pro Leu Leu Gly Phe Ile Gly Arg Leu Asp Gly Gln
610 615 620

Lys Gly Val Glu Ile Ile Ala Asp Ala Met Pro Trp Ile Val Ser Gln
625 630 635 640

Asp Val Gln Leu Val Met Leu Gly Thr Gly Arg His Asp Leu Glu Gly
645 650 655

Met Leu Arg His Phe Glu Arg Glu His His Asp Lys Val Arg Gly Trp
660 665 670

Val Gly Phe Ser Val Arg Leu Ala His Arg Ile Thr Ala Gly Ala Asp
675 680 685

Ala Leu Leu Met Pro Ser Arg Phe Glu Pro Cys Gly Leu Asn Gln Leu
690 695 700

Tyr Ala Met Ala Tyr Gly Thr Val Pro Val Val His Ala Val Gly Gly
705 710 715 720

Leu Arg Asp Thr Val Pro Pro Phe Asp Pro Phe Asn His Ser Gly Leu
725 730 735

Gly Trp Thr Phe Asp Arg Ala Glu Ala Gln Lys Leu Ile Glu Ala Leu
740 745 750

Gly His Cys Leu Arg Thr Tyr Arg Asp Tyr Lys Glu Ser Trp Arg Gly
755 760 765

Leu Gln Glu Arg Gly Met Ser Gln Asp Phe Ser Trp Glu His Ala Ala
770 775 780

Lys Leu Tyr Glu Asp Val Leu Val Lys Ala Lys Tyr Gln Trp
785 790 795

<210> 3
<211> 2800
<212> DNA
<213> Triticum aestivum

<220>
<221> CDS
<222> (89)..(2488)

<400> 3
gctgccacca cctccgcctg cgccgcgctc tgggcggagg accaacccgc gcatcgtacc 60

atcgcccgcc ccgatcccgg ccgccgcc atg tcg tcg gcg gtc gcg tcc gcc 112
Met Ser Ser Ala Val Ala Ser Ala
1 5

gcg tcc ttc ctc gcg ctc gcc tcc gcc tcc ccc ggg aga tca cgc agg 160
Ala Ser Phe Leu Ala Leu Ala Ser Ala Ser Pro Gly Arg Ser Arg Arg
10 15 20

cgg gcg agg gtg agc gcg ccg cca ccc cac gcc ggg gcc ggc agg ctg 208
Arg Ala Arg Val Ser Ala Pro Pro Pro His Ala Gly Ala Gly Arg Leu
25 30 35 40

cac tgg ccg ccg tgg ccg ccg cag cgc acg gct cgc gac gga ggt gtg 256
His Trp Pro Pro Trp Pro Pro Gln Arg Thr Ala Arg Asp Gly Gly Val
45 50 55

gcc gcg cgc gcc gcc ggg aag aag gac gcg agg gtc gac gac gac gcc 304
Ala Ala Arg Ala Ala Gly Lys Lys Asp Ala Arg Val Asp Asp Asp Ala
60 65 70

gcg tcc gcg agg cag ccc cgc gca cgc cgc ggt ggc gcc gcc acc aag 352
Ala Ser Ala Arg Gln Pro Arg Ala Arg Arg Gly Gly Ala Ala Thr Lys
75 80 85

gtc gcg gag cgg agg gat ccc gtc aag acg ctc gat cgc gac gcc gcg 400
Val Ala Glu Arg Arg Asp Pro Val Lys Thr Leu Asp Arg Asp Ala Ala
90 95 100

gaa ggt ggc gcg ccg gca ccg ccg gca ccg agg cag gac gcc gcc cgt 448
Glu Gly Gly Ala Pro Ala Pro Pro Ala Pro Arg Gln Asp Ala Ala Arg
105 110 115 120

cca ccg agt atg aac ggc acg ccg gtg aac ggt gag aac aaa tct acc 496
Pro Pro Ser Met Asn Gly Thr Pro Val Asn Gly Glu Asn Lys Ser Thr
125 130 135

ggc ggc ggc ggc gcg acc aaa gac agc ggg ctg ccc gca ccc gca cgc 544

```
Gly Gly Gly Gly Ala Thr Lys Asp Ser Gly Leu Pro Ala Pro Ala Arg
                140                 145                 150

gcg ccc cat ccg tcg acc cag aac aga gta cca gtg aac ggt gaa aac      592
Ala Pro His Pro Ser Thr Gln Asn Arg Val Pro Val Asn Gly Glu Asn
            155                 160                 165

aaa gct aac gtc gcc tcg ccg ccg acg agc ata gcc gag gtc gtg gct      640
Lys Ala Asn Val Ala Ser Pro Pro Thr Ser Ile Ala Glu Val Val Ala
        170                 175                 180

ccg gat tcc gca gct acc att tcc atc agt gac aag gcg ccg gag tcc      688
Pro Asp Ser Ala Ala Thr Ile Ser Ile Ser Asp Lys Ala Pro Glu Ser
185                 190                 195                 200

gtt gtc cca gcc gag aag ccg ccg ccg tcg tcc ggc tca aat ttc gtg      736
Val Val Pro Ala Glu Lys Pro Pro Pro Ser Ser Gly Ser Asn Phe Val
                205                 210                 215

gtc tcg gct tct gct ccc agg ctg gac att gac agc gat gtt gaa cct      784
Val Ser Ala Ser Ala Pro Arg Leu Asp Ile Asp Ser Asp Val Glu Pro
            220                 225                 230

gaa ctg aag aag ggt gcg gtc atc gtc gaa gaa gct cca aac cca aag      832
Glu Leu Lys Lys Gly Ala Val Ile Val Glu Glu Ala Pro Asn Pro Lys
            235                 240                 245

gct ctt tcg ccg cct gca gcc ccc gct gta caa gaa gac ctt tgg gac      880
Ala Leu Ser Pro Pro Ala Ala Pro Ala Val Gln Glu Asp Leu Trp Asp
        250                 255                 260

ttc aag aaa tac att ggc ttc gag gag ccc gtg gag gcc aag gat gat      928
Phe Lys Lys Tyr Ile Gly Phe Glu Glu Pro Val Glu Ala Lys Asp Asp
265                 270                 275                 280

ggc tgg gct gtt gca gat gat gcg ggc tcc ttt gaa cat cac cag aac      976
Gly Trp Ala Val Ala Asp Asp Ala Gly Ser Phe Glu His His Gln Asn
                285                 290                 295

cat gat tcc gga cct ttg gca ggg gag aac gtc atg aac gtg gtc gtc     1024
His Asp Ser Gly Pro Leu Ala Gly Glu Asn Val Met Asn Val Val Val
            300                 305                 310

gtg gct gct gaa tgt tct ccc tgg tgc aaa aca ggt ggt ctt gga gat     1072
Val Ala Ala Glu Cys Ser Pro Trp Cys Lys Thr Gly Gly Leu Gly Asp
            315                 320                 325

gtt gcc ggt gct ttg ccc aag gct ttg gcg aag aga gga cat cgt gtt     1120
Val Ala Gly Ala Leu Pro Lys Ala Leu Ala Lys Arg Gly His Arg Val
        330                 335                 340

atg gtt gtg gta cca agg tat ggg gac tat gag gaa gcc tac gat gtc     1168
Met Val Val Val Pro Arg Tyr Gly Asp Tyr Glu Glu Ala Tyr Asp Val
345                 350                 355                 360

gga gtc cga aaa tac tac aag gct gct gga cag gat atg gaa gtg aat     1216
Gly Val Arg Lys Tyr Tyr Lys Ala Ala Gly Gln Asp Met Glu Val Asn
                365                 370                 375

tat ttc cat gct tat atc gat gga gtt gat ttt gtg ttc att gac gct     1264
Tyr Phe His Ala Tyr Ile Asp Gly Val Asp Phe Val Phe Ile Asp Ala
            380                 385                 390

cct ctc ttc cga cac cgc cag gaa gac att tat ggg ggc agc aga cag     1312
Pro Leu Phe Arg His Arg Gln Glu Asp Ile Tyr Gly Gly Ser Arg Gln
```

```
        395                 400                 405

gaa att atg aag cgc atg att ttg ttc tgc aag gcc gct gtc gag gtt    1360
Glu Ile Met Lys Arg Met Ile Leu Phe Cys Lys Ala Ala Val Glu Val
    410                 415                 420

cct tgg cac gtt cca tgc ggc ggt gtc cct tat ggg gat gga aat ctg    1408
Pro Trp His Val Pro Cys Gly Gly Val Pro Tyr Gly Asp Gly Asn Leu
425                 430                 435                 440

gtg ttt att gca aat gat tgg cac acg gca ctc ctg cct gtc tat ctg    1456
Val Phe Ile Ala Asn Asp Trp His Thr Ala Leu Leu Pro Val Tyr Leu
                445                 450                 455

aaa gca tat tac agg gac cat ggt ttg atg cag tac act cgg tcc att    1504
Lys Ala Tyr Tyr Arg Asp His Gly Leu Met Gln Tyr Thr Arg Ser Ile
            460                 465                 470

atg gtg ata cat aac atc gcg cac cag ggc cgt ggc cca gta gat gaa    1552
Met Val Ile His Asn Ile Ala His Gln Gly Arg Gly Pro Val Asp Glu
        475                 480                 485

ttc ccg ttc acc gag ttg cct gag cac tac ctg gaa cac ttc aga ctg    1600
Phe Pro Phe Thr Glu Leu Pro Glu His Tyr Leu Glu His Phe Arg Leu
    490                 495                 500

tac gac ccc gtg ggt ggt gag cac gcc aac tac ttc gcc gcc ggc ctg    1648
Tyr Asp Pro Val Gly Gly Glu His Ala Asn Tyr Phe Ala Ala Gly Leu
505                 510                 515                 520

aag atg gcg gac cag gtt gtc gtg gtg agc ccc ggg tac ctg tgg gag    1696
Lys Met Ala Asp Gln Val Val Val Val Ser Pro Gly Tyr Leu Trp Glu
                525                 530                 535

ctc aag acg gtg gag ggc ggc tgg ggg ctt cac gac atc ata cgg cag    1744
Leu Lys Thr Val Glu Gly Gly Trp Gly Leu His Asp Ile Ile Arg Gln
            540                 545                 550

aac gac tgg aag acc cgc ggc atc gtc aac ggc atc gac aac atg gag    1792
Asn Asp Trp Lys Thr Arg Gly Ile Val Asn Gly Ile Asp Asn Met Glu
        555                 560                 565

tgg aac ccc gag gtg gac gtc cac ctc aag tcg gac ggc tac acc aac    1840
Trp Asn Pro Glu Val Asp Val His Leu Lys Ser Asp Gly Tyr Thr Asn
    570                 575                 580

ttc tcc ctg ggg acg ctg gac tcc ggc aag cgg cag tgc aag gag gcc    1888
Phe Ser Leu Gly Thr Leu Asp Ser Gly Lys Arg Gln Cys Lys Glu Ala
585                 590                 595                 600

ctg cag cgc gag ctg ggc ctg cag gtc cgc gcc gac gtg ccg ctg ctc    1936
Leu Gln Arg Glu Leu Gly Leu Gln Val Arg Ala Asp Val Pro Leu Leu
                605                 610                 615

ggc ttc atc ggc cgc ctg gac ggg cag aag ggc gtg gag atc atc gcg    1984
Gly Phe Ile Gly Arg Leu Asp Gly Gln Lys Gly Val Glu Ile Ile Ala
            620                 625                 630

gac gcc atg ccc tgg atc gtg agc cag gac gtg cag ctg gtc atg ctg    2032
Asp Ala Met Pro Trp Ile Val Ser Gln Asp Val Gln Leu Val Met Leu
        635                 640                 645

ggc acc ggc cgc cac gac ctg gag agc atg ctg cgg cac ttc gag cgg    2080
Gly Thr Gly Arg His Asp Leu Glu Ser Met Leu Arg His Phe Glu Arg
    650                 655                 660
```

gag cac cac gac aag gtg cgc ggg tgg gtg ggg ttc tcc gtg cgc ctg 2128
Glu His His Asp Lys Val Arg Gly Trp Val Gly Phe Ser Val Arg Leu
665 670 675 680

gcg cac cgg atc acg gcg ggc gcc gac gcg ctc ctc atg ccc tcc cgg 2176
Ala His Arg Ile Thr Ala Gly Ala Asp Ala Leu Leu Met Pro Ser Arg
685 690 695

ttc gag ccg tgc ggg ttg aac cag ctt tac gcc atg gcc tac ggc acc 2224
Phe Glu Pro Cys Gly Leu Asn Gln Leu Tyr Ala Met Ala Tyr Gly Thr
700 705 710

gtc ccc gtc gtg cac gcc gtc ggc ggg gtg agg gac acc gtg ccg ccg 2272
Val Pro Val Val His Ala Val Gly Gly Val Arg Asp Thr Val Pro Pro
715 720 725

ttc gac ccc ttc aac cac tcc ggc ctc ggg tgg acg ttc gac cgc gcc 2320
Phe Asp Pro Phe Asn His Ser Gly Leu Gly Trp Thr Phe Asp Arg Ala
730 735 740

gag gcg cac aag ctg atc gag gcg ctc ggg cac tgc ctc cgc acc tac 2368
Glu Ala His Lys Leu Ile Glu Ala Leu Gly His Cys Leu Arg Thr Tyr
745 750 755 760

cgg gac tac aag gag agc tgg agg ggc ctc cag gag cgc ggc atg tcg 2416
Arg Asp Tyr Lys Glu Ser Trp Arg Gly Leu Gln Glu Arg Gly Met Ser
765 770 775

cag gac ttc agc tgg gag cat gcc gcc aag ctc tac gag gac gtc ctc 2464
Gln Asp Phe Ser Trp Glu His Ala Ala Lys Leu Tyr Glu Asp Val Leu
780 785 790

ctc aag gcc aag tac cag tgg tga acgctagctg ctagccgctc cagccccgca 2518
Leu Lys Ala Lys Tyr Gln Trp
795 800

tgcgtgcatg catgagaggg tggaactgcg cattgcgccc gcaggaacgt gccatccttc 2578

tcgatgggag cgccggcatc cgcgaggtgc agtgacatga gaggtgtgtg tggttgagac 2638

gctgattccg atctcgatct ggtccgtagc agagtagagc ggacgtaggg aagcgctcct 2698

tgttgcaggt atatgggaat gttgtcaact tggtattgta gtttgctatg ttgtatgcgt 2758

tattacaatg ttgttactta ttcttgttaa aaaaaaaaaa aa 2800

<210> 4
<211> 799
<212> PRT
<213> Triticum aestivum

<400> 4
Met Ser Ser Ala Val Ala Ser Ala Ala Ser Phe Leu Ala Leu Ala Ser
1 5 10 15

Ala Ser Pro Gly Arg Ser Arg Arg Arg Ala Arg Val Ser Ala Pro Pro
20 25 30

Pro His Ala Gly Ala Gly Arg Leu His Trp Pro Pro Trp Pro Pro Gln
35 40 45

Arg Thr Ala Arg Asp Gly Gly Val Ala Ala Arg Ala Ala Gly Lys Lys
50 55 60

Asp Ala Arg Val Asp Asp Asp Ala Ala Ser Ala Arg Gln Pro Arg Ala
65 70 75 80

Arg Arg Gly Gly Ala Ala Thr Lys Val Ala Glu Arg Arg Asp Pro Val
85 90 95

Lys Thr Leu Asp Arg Asp Ala Ala Glu Gly Gly Ala Pro Ala Pro Pro
100 105 110

Ala Pro Arg Gln Asp Ala Ala Arg Pro Pro Ser Met Asn Gly Thr Pro
115 120 125

Val Asn Gly Glu Asn Lys Ser Thr Gly Gly Gly Gly Ala Thr Lys Asp
130 135 140

Ser Gly Leu Pro Ala Pro Ala Arg Ala Pro His Pro Ser Thr Gln Asn
145 150 155 160

Arg Val Pro Val Asn Gly Glu Asn Lys Ala Asn Val Ala Ser Pro Pro
165 170 175

Thr Ser Ile Ala Glu Val Val Ala Pro Asp Ser Ala Ala Thr Ile Ser
180 185 190

Ile Ser Asp Lys Ala Pro Glu Ser Val Val Pro Ala Glu Lys Pro Pro
195 200 205

Pro Ser Ser Gly Ser Asn Phe Val Val Ser Ala Ser Ala Pro Arg Leu
210 215 220

Asp Ile Asp Ser Asp Val Glu Pro Glu Leu Lys Lys Gly Ala Val Ile
225 230 235 240

Val Glu Glu Ala Pro Asn Pro Lys Ala Leu Ser Pro Pro Ala Ala Pro
245 250 255

Ala Val Gln Glu Asp Leu Trp Asp Phe Lys Lys Tyr Ile Gly Phe Glu
260 265 270

Glu Pro Val Glu Ala Lys Asp Asp Gly Trp Ala Val Ala Asp Asp Ala
275 280 285

Gly Ser Phe Glu His His Gln Asn His Asp Ser Gly Pro Leu Ala Gly
290 295 300

Glu Asn Val Met Asn Val Val Val Val Ala Ala Glu Cys Ser Pro Trp
305 310 315 320

Cys Lys Thr Gly Gly Leu Gly Asp Val Ala Gly Ala Leu Pro Lys Ala
325 330 335

Leu Ala Lys Arg Gly His Arg Val Met Val Val Val Pro Arg Tyr Gly
340 345 350

Asp Tyr Glu Glu Ala Tyr Asp Val Gly Val Arg Lys Tyr Tyr Lys Ala
355 360 365

Ala Gly Gln Asp Met Glu Val Asn Tyr Phe His Ala Tyr Ile Asp Gly
370 375 380

Val Asp Phe Val Phe Ile Asp Ala Pro Leu Phe Arg His Arg Gln Glu
385 390 395 400

Asp Ile Tyr Gly Gly Ser Arg Gln Glu Ile Met Lys Arg Met Ile Leu

405 410 415

Phe Cys Lys Ala Ala Val Glu Val Pro Trp His Val Pro Cys Gly Gly
420 425 430

Val Pro Tyr Gly Asp Gly Asn Leu Val Phe Ile Ala Asn Asp Trp His
435 440 445

Thr Ala Leu Leu Pro Val Tyr Leu Lys Ala Tyr Tyr Arg Asp His Gly
450 455 460

Leu Met Gln Tyr Thr Arg Ser Ile Met Val Ile His Asn Ile Ala His
465 470 475 480

Gln Gly Arg Gly Pro Val Asp Glu Phe Pro Phe Thr Glu Leu Pro Glu
485 490 495

His Tyr Leu Glu His Phe Arg Leu Tyr Asp Pro Val Gly Gly Glu His
500 505 510

Ala Asn Tyr Phe Ala Ala Gly Leu Lys Met Ala Asp Gln Val Val Val
515 520 525

Val Ser Pro Gly Tyr Leu Trp Glu Leu Lys Thr Val Glu Gly Gly Trp
530 535 540

Gly Leu His Asp Ile Ile Arg Gln Asn Asp Trp Lys Thr Arg Gly Ile
545 550 555 560

Val Asn Gly Ile Asp Asn Met Glu Trp Asn Pro Glu Val Asp Val His
565 570 575

Leu Lys Ser Asp Gly Tyr Thr Asn Phe Ser Leu Gly Thr Leu Asp Ser
580 585 590

Gly Lys Arg Gln Cys Lys Glu Ala Leu Gln Arg Glu Leu Gly Leu Gln
595 600 605

Val Arg Ala Asp Val Pro Leu Leu Gly Phe Ile Gly Arg Leu Asp Gly
610 615 620

Gln Lys Gly Val Glu Ile Ile Ala Asp Ala Met Pro Trp Ile Val Ser
625 630 635 640

Gln Asp Val Gln Leu Val Met Leu Gly Thr Gly Arg His Asp Leu Glu
645 650 655

Ser Met Leu Arg His Phe Glu Arg Glu His His Asp Lys Val Arg Gly
660 665 670

Trp Val Gly Phe Ser Val Arg Leu Ala His Arg Ile Thr Ala Gly Ala
675 680 685

Asp Ala Leu Leu Met Pro Ser Arg Phe Glu Pro Cys Gly Leu Asn Gln
690 695 700

Leu Tyr Ala Met Ala Tyr Gly Thr Val Pro Val Val His Ala Val Gly
705 710 715 720

Gly Val Arg Asp Thr Val Pro Pro Phe Asp Pro Phe Asn His Ser Gly
725 730 735

Leu Gly Trp Thr Phe Asp Arg Ala Glu Ala His Lys Leu Ile Glu Ala
740 745 750

Leu Gly His Cys Leu Arg Thr Tyr Arg Asp Tyr Lys Glu Ser Trp Arg
755 760 765

Gly Leu Gln Glu Arg Gly Met Ser Gln Asp Phe Ser Trp Glu His Ala
770 775 780

Ala Lys Leu Tyr Glu Asp Val Leu Leu Lys Ala Lys Tyr Gln Trp
785 790 795

<210> 5
<211> 2107
<212> DNA
<213> Triticum aestivum

<220>
<221> CDS
<222> (1)..(1794)

<400> 5
cca gct gag aag acg ccg ccg tcg tcc ggc tca aat ttc gag tcc tcg 48
Pro Ala Glu Lys Thr Pro Pro Ser Ser Gly Ser Asn Phe Glu Ser Ser
1 5 10 15

gcc tct gct ccc ggg tct gac act gtc agc gac gtg gaa caa gaa ctg 96
Ala Ser Ala Pro Gly Ser Asp Thr Val Ser Asp Val Glu Gln Glu Leu
20 25 30

aag aag ggt gcg gtc gtt gtc gaa gaa gct cca aag cca aag gct ctt 144
Lys Lys Gly Ala Val Val Val Glu Glu Ala Pro Lys Pro Lys Ala Leu
35 40 45

tcg ccg cct gca gcc ccc gct gta caa gaa gac ctt tgg gat ttc aag 192
Ser Pro Pro Ala Ala Pro Ala Val Gln Glu Asp Leu Trp Asp Phe Lys
50 55 60

aaa tac att ggt ttc gag gag ccc gtg gag gcc aag gat gat ggc cgg 240
Lys Tyr Ile Gly Phe Glu Glu Pro Val Glu Ala Lys Asp Asp Gly Arg
65 70 75 80

gct gtc gca gat gat gcg ggc tcc ttt gaa cac cac cag aat cac gac 288
Ala Val Ala Asp Asp Ala Gly Ser Phe Glu His His Gln Asn His Asp
85 90 95

tcc gga cct ttg gca ggg gag aat gtc atg aac gtg gtc gtc gtg gct 336
Ser Gly Pro Leu Ala Gly Glu Asn Val Met Asn Val Val Val Val Ala
100 105 110

gct gag tgt tct ccc tgg tgc aaa aca ggt ggt ctg gga gat gtt gcg 384
Ala Glu Cys Ser Pro Trp Cys Lys Thr Gly Gly Leu Gly Asp Val Ala
115 120 125

ggt gct ctg ccc aag gct ttg gca aag aga gga cat cgt gtt atg gtt 432
Gly Ala Leu Pro Lys Ala Leu Ala Lys Arg Gly His Arg Val Met Val
130 135 140

gtg gta cca agg tat ggg gac tat gaa gaa cct acg gat gtc gga gtc 480
Val Val Pro Arg Tyr Gly Asp Tyr Glu Glu Pro Thr Asp Val Gly Val
145 150 155 160

cga aaa tac tac aag gct gct gga cag gat atg gaa gtg aat tat ttc 528
Arg Lys Tyr Tyr Lys Ala Ala Gly Gln Asp Met Glu Val Asn Tyr Phe
165 170 175

cat gct tat atc gat gga gtt gat ttt gtg ttc att gac gct cct ctc 576

```
His Ala Tyr Ile Asp Gly Val Asp Phe Val Phe Ile Asp Ala Pro Leu
            180                 185                 190

ttc cga cac cga gag gaa gac att tat ggg ggc agc aga cag gaa att     624
Phe Arg His Arg Glu Glu Asp Ile Tyr Gly Gly Ser Arg Gln Glu Ile
        195                 200                 205

atg aag cgc atg att ttg ttc tgc aag gcc gct gtt gag gtt cca tgg     672
Met Lys Arg Met Ile Leu Phe Cys Lys Ala Ala Val Glu Val Pro Trp
    210                 215                 220

cac gtt cca tgc ggc ggt gtc cct tat ggg gat gga aat ctg gtg ttt     720
His Val Pro Cys Gly Gly Val Pro Tyr Gly Asp Gly Asn Leu Val Phe
225                 230                 235                 240

att gca aat gat tgg cac acg gca ctc ctg cct gtc tat ctg aaa gca     768
Ile Ala Asn Asp Trp His Thr Ala Leu Leu Pro Val Tyr Leu Lys Ala
                245                 250                 255

tat tac agg gac cat ggt ttg atg cag tac act cgg tcc att atg gtg     816
Tyr Tyr Arg Asp His Gly Leu Met Gln Tyr Thr Arg Ser Ile Met Val
            260                 265                 270

ata cat aac atc gct cac cag ggc cgt ggc cct gta gat gaa ttc ccg     864
Ile His Asn Ile Ala His Gln Gly Arg Gly Pro Val Asp Glu Phe Pro
        275                 280                 285

ttc acc gag ttg cct gag cac tac ctg gaa cac ttc aga ctg tac gac     912
Phe Thr Glu Leu Pro Glu His Tyr Leu Glu His Phe Arg Leu Tyr Asp
    290                 295                 300

ccc gtg ggt ggt gaa cac gcc aac tac ttc gcc gcc ggc ctg aag atg     960
Pro Val Gly Gly Glu His Ala Asn Tyr Phe Ala Ala Gly Leu Lys Met
305                 310                 315                 320

gcg gac cag gtt gtc gtg gtg agc ccc ggg tac ctg tgg gag ctg aag     1008
Ala Asp Gln Val Val Val Val Ser Pro Gly Tyr Leu Trp Glu Leu Lys
                325                 330                 335

acg gtg gag ggc ggc tgg ggg ctt cac gac atc ata cgg cag aac gac     1056
Thr Val Glu Gly Gly Trp Gly Leu His Asp Ile Ile Arg Gln Asn Asp
            340                 345                 350

tgg aag acc cgc ggc atc gtc aac ggc atc gac aac atg gag tgg aac     1104
Trp Lys Thr Arg Gly Ile Val Asn Gly Ile Asp Asn Met Glu Trp Asn
        355                 360                 365

ccc gag gtg gac gcc cac ctc aag tcg gac ggc tac acc aac ttc tcc     1152
Pro Glu Val Asp Ala His Leu Lys Ser Asp Gly Tyr Thr Asn Phe Ser
    370                 375                 380

ctg agg acg ctg gac tcc ggc aag cgg cag tgc aag gag gcc ctg cag     1200
Leu Arg Thr Leu Asp Ser Gly Lys Arg Gln Cys Lys Glu Ala Leu Gln
385                 390                 395                 400

cgc gag ctg ggc ctg cag gtc cgc gcc gac gtg ccg ctg ctc ggc ttc     1248
Arg Glu Leu Gly Leu Gln Val Arg Ala Asp Val Pro Leu Leu Gly Phe
                405                 410                 415

atc ggc cgc ctg gac ggg cag aag ggc gtg gag atc atc gcg gac gcc     1296
Ile Gly Arg Leu Asp Gly Gln Lys Gly Val Glu Ile Ile Ala Asp Ala
            420                 425                 430

atg ccc tgg atc gtg agc cag gac gtg cag ctg gtg atg ctg ggc acc     1344
Met Pro Trp Ile Val Ser Gln Asp Val Gln Leu Val Met Leu Gly Thr
```

435 440 445

ggg cgc cac gac ctg gag agc atg ctg cag cac ttc gag cgg gag cac 1392
Gly Arg His Asp Leu Glu Ser Met Leu Gln His Phe Glu Arg Glu His
450 455 460

cac gac aag gtg cgc ggg tgg gtg ggg ttc tcc gtg cgc ctg gcg cac 1440
His Asp Lys Val Arg Gly Trp Val Gly Phe Ser Val Arg Leu Ala His
465 470 475 480

cgg atc acg gcg ggg gcg gac gcg ctc ctc atg ccc tcc cgg ttc gtg 1488
Arg Ile Thr Ala Gly Ala Asp Ala Leu Leu Met Pro Ser Arg Phe Val
485 490 495

ccg tgc ggg ctg aac cag ctc tac gcc atg gcc tac ggc acc gtc ccc 1536
Pro Cys Gly Leu Asn Gln Leu Tyr Ala Met Ala Tyr Gly Thr Val Pro
500 505 510

gtc gtg cac gcc gtc ggc ggc ctc agg gac acc gtg ccg ccg ttc gac 1584
Val Val His Ala Val Gly Gly Leu Arg Asp Thr Val Pro Pro Phe Asp
515 520 525

ccc ttc aac cac tcc ggg ctc ggg tgg acg ttc gac cgc gcc gag gcg 1632
Pro Phe Asn His Ser Gly Leu Gly Trp Thr Phe Asp Arg Ala Glu Ala
530 535 540

cac aag ctg atc gag gcg ctc ggg cac tgc ctc cgc acc tac cga gac 1680
His Lys Leu Ile Glu Ala Leu Gly His Cys Leu Arg Thr Tyr Arg Asp
545 550 555 560

ttc aag gag agc tgg agg gcc ctc cag gag cgc ggc atg tcg cag gac 1728
Phe Lys Glu Ser Trp Arg Ala Leu Gln Glu Arg Gly Met Ser Gln Asp
565 570 575

ttc agc tgg gag cac gcc gcc aag ctc tac gag gac gtc ctc gtc aag 1776
Phe Ser Trp Glu His Ala Ala Lys Leu Tyr Glu Asp Val Leu Val Lys
580 585 590

gcc aag tac cag tgg tga acgctagctg ctagccgctc cagccccgca 1824
Ala Lys Tyr Gln Trp
595

tgcgtgcatg acaggatgga actgcattgc gcacgcagga aagtgccatg gagcgccggc 1884

atccgcgaag tacagtgaca tgaggtgtgt gtggttgaga cgctgattcc aatccggccc 1944

gtagcagagt agagcggagg tatatgggaa tcttaacttg gtattgtaat ttgttatgtt 2004

gtgtgcatta ttacaatgtt gttacttatt cttgttaagt cggaggccaa gggcgaaagc 2064

tagctcacat gtctgatgga tgcaaaaaaa aaaaaaaaaa aaa 2107

<210> 6
<211> 597
<212> PRT
<213> Triticum aestivum

<400> 6
Pro Ala Glu Lys Thr Pro Pro Ser Ser Gly Ser Asn Phe Glu Ser Ser
1 5 10 15

Ala Ser Ala Pro Gly Ser Asp Thr Val Ser Asp Val Glu Gln Glu Leu
20 25 30

```
Lys Lys Gly Ala Val Val Val Glu Glu Ala Pro Lys Pro Lys Ala Leu
         35                  40                  45

Ser Pro Pro Ala Ala Pro Ala Val Gln Glu Asp Leu Trp Asp Phe Lys
     50                  55                  60

Lys Tyr Ile Gly Phe Glu Glu Pro Val Glu Ala Lys Asp Asp Gly Arg
 65                  70                  75                  80

Ala Val Ala Asp Asp Ala Gly Ser Phe Glu His His Gln Asn His Asp
                 85                  90                  95

Ser Gly Pro Leu Ala Gly Glu Asn Val Met Asn Val Val Val Val Ala
            100                 105                 110

Ala Glu Cys Ser Pro Trp Cys Lys Thr Gly Gly Leu Gly Asp Val Ala
        115                 120                 125

Gly Ala Leu Pro Lys Ala Leu Ala Lys Arg Gly His Arg Val Met Val
    130                 135                 140

Val Val Pro Arg Tyr Gly Asp Tyr Glu Glu Pro Thr Asp Val Gly Val
145                 150                 155                 160

Arg Lys Tyr Tyr Lys Ala Ala Gly Gln Asp Met Glu Val Asn Tyr Phe
                165                 170                 175

His Ala Tyr Ile Asp Gly Val Asp Phe Val Phe Ile Asp Ala Pro Leu
            180                 185                 190

Phe Arg His Arg Glu Glu Asp Ile Tyr Gly Gly Ser Arg Gln Glu Ile
        195                 200                 205

Met Lys Arg Met Ile Leu Phe Cys Lys Ala Ala Val Glu Val Pro Trp
    210                 215                 220

His Val Pro Cys Gly Gly Val Pro Tyr Gly Asp Gly Asn Leu Val Phe
225                 230                 235                 240

Ile Ala Asn Asp Trp His Thr Ala Leu Leu Pro Val Tyr Leu Lys Ala
                245                 250                 255

Tyr Tyr Arg Asp His Gly Leu Met Gln Tyr Thr Arg Ser Ile Met Val
            260                 265                 270

Ile His Asn Ile Ala His Gln Gly Arg Gly Pro Val Asp Glu Phe Pro
        275                 280                 285

Phe Thr Glu Leu Pro Glu His Tyr Leu Glu His Phe Arg Leu Tyr Asp
    290                 295                 300

Pro Val Gly Gly Glu His Ala Asn Tyr Phe Ala Ala Gly Leu Lys Met
305                 310                 315                 320

Ala Asp Gln Val Val Val Val Ser Pro Gly Tyr Leu Trp Glu Leu Lys
                325                 330                 335

Thr Val Glu Gly Gly Trp Gly Leu His Asp Ile Ile Arg Gln Asn Asp
            340                 345                 350

Trp Lys Thr Arg Gly Ile Val Asn Gly Ile Asp Asn Met Glu Trp Asn
        355                 360                 365

Pro Glu Val Asp Ala His Leu Lys Ser Asp Gly Tyr Thr Asn Phe Ser
    370                 375                 380
```

```
Leu Arg Thr Leu Asp Ser Gly Lys Arg Gln Cys Lys Glu Ala Leu Gln
385                 390                 395                 400

Arg Glu Leu Gly Leu Gln Val Arg Ala Asp Val Pro Leu Leu Gly Phe
                405                 410                 415

Ile Gly Arg Leu Asp Gly Gln Lys Gly Val Glu Ile Ile Ala Asp Ala
            420                 425                 430

Met Pro Trp Ile Val Ser Gln Asp Val Gln Leu Val Met Leu Gly Thr
        435                 440                 445

Gly Arg His Asp Leu Glu Ser Met Leu Gln His Phe Glu Arg Glu His
    450                 455                 460

His Asp Lys Val Arg Gly Trp Val Gly Phe Ser Val Arg Leu Ala His
465                 470                 475                 480

Arg Ile Thr Ala Gly Ala Asp Ala Leu Leu Met Pro Ser Arg Phe Val
                485                 490                 495

Pro Cys Gly Leu Asn Gln Leu Tyr Ala Met Ala Tyr Gly Thr Val Pro
            500                 505                 510

Val Val His Ala Val Gly Gly Leu Arg Asp Thr Val Pro Pro Phe Asp
        515                 520                 525

Pro Phe Asn His Ser Gly Leu Gly Trp Thr Phe Asp Arg Ala Glu Ala
    530                 535                 540

His Lys Leu Ile Glu Ala Leu Gly His Cys Leu Arg Thr Tyr Arg Asp
545                 550                 555                 560

Phe Lys Glu Ser Trp Arg Ala Leu Gln Glu Arg Gly Met Ser Gln Asp
                565                 570                 575

Phe Ser Trp Glu His Ala Ala Lys Leu Tyr Glu Asp Val Leu Val Lys
            580                 585                 590

Ala Lys Tyr Gln Trp
        595


<210> 7
<211> 5352
<212> DNA
<213> Triticum aestivum

<220>
<221> CDS
<222> (29)..(4921)

<400> 7
cggcacgagg tttagtaggt tccgggaa atg gag atg tct ctc tgg cca cgg      52
                               Met Glu Met Ser Leu Trp Pro Arg
                               1               5

agc ccc ctg tgc cct cgg agc agg cag ccg ctc gtc gtc gtc cgg ccg     100
Ser Pro Leu Cys Pro Arg Ser Arg Gln Pro Leu Val Val Val Arg Pro
    10                  15                  20

gcc ggc cgc ggc ggc ctc acg cag cct ttt ttg atg aat ggc aga ttt     148
Ala Gly Arg Gly Gly Leu Thr Gln Pro Phe Leu Met Asn Gly Arg Phe
25                  30                  35                  40
```

act cga agc agg acc ctt cga tgc atg gta gca agt tca gat cct cct 196
Thr Arg Ser Arg Thr Leu Arg Cys Met Val Ala Ser Ser Asp Pro Pro
45 50 55

aat agg aaa tca aga agg atg gta cca cct cag gtt aaa gtc att tct 244
Asn Arg Lys Ser Arg Arg Met Val Pro Pro Gln Val Lys Val Ile Ser
60 65 70

tct aga gga tat acg aca aga ctc att gtt gaa cca agc aac gag aat 292
Ser Arg Gly Tyr Thr Thr Arg Leu Ile Val Glu Pro Ser Asn Glu Asn
75 80 85

aca gaa cac aat aat cgg gat gaa gaa act ctt gat aca tac aat gcg 340
Thr Glu His Asn Asn Arg Asp Glu Glu Thr Leu Asp Thr Tyr Asn Ala
90 95 100

cta tta agt acc gag aca gca gaa tgg aca gat aat aga gaa gcc gag 388
Leu Leu Ser Thr Glu Thr Ala Glu Trp Thr Asp Asn Arg Glu Ala Glu
105 110 115 120

act gct aaa gcg gac tcg tcg caa aat gct tta agc agt tct ata att 436
Thr Ala Lys Ala Asp Ser Ser Gln Asn Ala Leu Ser Ser Ser Ile Ile
125 130 135

ggg gaa gtg gat gtg gcg gat gaa gat ata ctt gcg gct gat ctg aca 484
Gly Glu Val Asp Val Ala Asp Glu Asp Ile Leu Ala Ala Asp Leu Thr
140 145 150

gtg tat tca ttg agc agt gta atg aag aag gaa gtg gat gca gcg gac 532
Val Tyr Ser Leu Ser Ser Val Met Lys Lys Glu Val Asp Ala Ala Asp
155 160 165

aaa gct aga gtt aaa gaa gac gca ttt gag ctg gat ttn gcc agc act 580
Lys Ala Arg Val Lys Glu Asp Ala Phe Glu Leu Asp Xaa Ala Ser Thr
170 175 180

aca ttg aga agt gtg ata gta gat gtg atg gat cat aan tgg gac tgt 628
Thr Leu Arg Ser Val Ile Val Asp Val Met Asp His Xaa Trp Asp Cys
185 190 195 200

caa gag aca ttg aga agt gtg ata gta gat gtg atg gat cat aat ggg 676
Gln Glu Thr Leu Arg Ser Val Ile Val Asp Val Met Asp His Asn Gly
205 210 215

act gta caa gag aca ttg aga agt gtg ata gta gat gtg atg gat gat 724
Thr Val Gln Glu Thr Leu Arg Ser Val Ile Val Asp Val Met Asp Asp
220 225 230

gcg gcg gac aaa gct aga gtt gaa gaa gac gta ttt gag ctg gat ttg 772
Ala Ala Asp Lys Ala Arg Val Glu Glu Asp Val Phe Glu Leu Asp Leu
235 240 245

tca gga aat att tca agc agt gcg acg acc gtg gaa cta gat gcg gtt 820
Ser Gly Asn Ile Ser Ser Ser Ala Thr Thr Val Glu Leu Asp Ala Val
250 255 260

gac gaa gtc ggg cct gtt caa gac aaa ttt gag gcg acc tca tca gga 868
Asp Glu Val Gly Pro Val Gln Asp Lys Phe Glu Ala Thr Ser Ser Gly
265 270 275 280

aat gtt tca aac agt gca acg gta cgg gaa gtg gat gca agt gat gaa 916
Asn Val Ser Asn Ser Ala Thr Val Arg Glu Val Asp Ala Ser Asp Glu
285 290 295

```
gct ggg aat gat caa ggc ata ttt aga gca gat ttg tca gga aat gtt     964
Ala Gly Asn Asp Gln Gly Ile Phe Arg Ala Asp Leu Ser Gly Asn Val
            300                 305                 310

ttt tca agc agt aca aca gtg gaa gtg ggt gca gtg gat gaa gct ggg     1012
Phe Ser Ser Ser Thr Thr Val Glu Val Gly Ala Val Asp Glu Ala Gly
        315                 320                 325

tct ata aag gac agg ttt gag acg gat tcg tca gga aat gtt tca aca     1060
Ser Ile Lys Asp Arg Phe Glu Thr Asp Ser Ser Gly Asn Val Ser Thr
    330                 335                 340

agt gcg ccg atg tgg gat gca att gat gaa acc gtg gct gat caa gac     1108
Ser Ala Pro Met Trp Asp Ala Ile Asp Glu Thr Val Ala Asp Gln Asp
345                 350                 355                 360

aca ttt gag gcg gat ttg tcg gga aat gct tca agc tgc gca aca tac     1156
Thr Phe Glu Ala Asp Leu Ser Gly Asn Ala Ser Ser Cys Ala Thr Tyr
                365                 370                 375

aga gaa gtg gat gat gtg gtg gat gaa act aga tca gaa gag gaa aca     1204
Arg Glu Val Asp Asp Val Val Asp Glu Thr Arg Ser Glu Glu Glu Thr
            380                 385                 390

ttt gca atg gat ttg ttt gca agt gaa tca ggc cat gag aaa cat atg     1252
Phe Ala Met Asp Leu Phe Ala Ser Glu Ser Gly His Glu Lys His Met
        395                 400                 405

gca gtg gat tat gtg ggt gaa gct acc gat gaa gaa gag act tac caa     1300
Ala Val Asp Tyr Val Gly Glu Ala Thr Asp Glu Glu Glu Thr Tyr Gln
    410                 415                 420

cag caa tat cca gta ccg tct tca ttc tct atg tgg gac aag gct att     1348
Gln Gln Tyr Pro Val Pro Ser Ser Phe Ser Met Trp Asp Lys Ala Ile
425                 430                 435                 440

gct aaa aca ggt gta agt ttg aat cct gag ctg cga ctt gtc agg gtt     1396
Ala Lys Thr Gly Val Ser Leu Asn Pro Glu Leu Arg Leu Val Arg Val
                445                 450                 455

gaa gaa caa ggc aaa gta aat ttt agt gat aaa aaa gac ctg tca att     1444
Glu Glu Gln Gly Lys Val Asn Phe Ser Asp Lys Lys Asp Leu Ser Ile
            460                 465                 470

gat gat tta cca gga caa aac caa tcg atc att ggt tcc tat aaa caa     1492
Asp Asp Leu Pro Gly Gln Asn Gln Ser Ile Ile Gly Ser Tyr Lys Gln
        475                 480                 485

gat aaa tca att gct gat gtt gcg gga ccg acc caa tca att ttt ggt     1540
Asp Lys Ser Ile Ala Asp Val Ala Gly Pro Thr Gln Ser Ile Phe Gly
    490                 495                 500

tct agt aaa caa cac cgg tca att gtt gct ttc ccc aaa caa aac cag     1588
Ser Ser Lys Gln His Arg Ser Ile Val Ala Phe Pro Lys Gln Asn Gln
505                 510                 515                 520

tca att gtt agt gtc act gag caa aag cag tcc ata gtt gga ttc cgt     1636
Ser Ile Val Ser Val Thr Glu Gln Lys Gln Ser Ile Val Gly Phe Arg
                525                 530                 535

agt caa gat ctt tcg gct gtt agt ctc cct aaa caa aac gta cca att     1684
Ser Gln Asp Leu Ser Ala Val Ser Leu Pro Lys Gln Asn Val Pro Ile
            540                 545                 550

gtt ggg tac gtc gag aga ggg tca aac naa aag caa gtt cct gtt gtt     1732
```

Val Gly Tyr Val Glu Arg Gly Ser Asn Xaa Lys Gln Val Pro Val Val
555 560 565

gat aga cag gat gca ttg tat gtg aat gga ctg gaa gct aag gag gga 1780
Asp Arg Gln Asp Ala Leu Tyr Val Asn Gly Leu Glu Ala Lys Glu Gly
570 575 580

gat cac aca tcc gag aaa act gat gag gat gcg ctt cat gta aag ttt 1828
Asp His Thr Ser Glu Lys Thr Asp Glu Asp Ala Leu His Val Lys Phe
585 590 595 600

aat gtt gac aat gtg ttg cgg aag cat cag gca gat aga acc caa gca 1876
Asn Val Asp Asn Val Leu Arg Lys His Gln Ala Asp Arg Thr Gln Ala
605 610 615

gtg gaa aag aaa act tgg aag aaa gtt gat gag gaa cat ctt tac atg 1924
Val Glu Lys Lys Thr Trp Lys Lys Val Asp Glu Glu His Leu Tyr Met
620 625 630

act gaa cat cag aaa cgt gct gcc gaa gga cag atg gta gtt aac gag 1972
Thr Glu His Gln Lys Arg Ala Ala Glu Gly Gln Met Val Val Asn Glu
635 640 645

gat gag ctt tct ata act gaa att gga atg ggg aga ggt gat aaa att 2020
Asp Glu Leu Ser Ile Thr Glu Ile Gly Met Gly Arg Gly Asp Lys Ile
650 655 660

cag cat gtg ctt tct gag gaa gag ctt tca tgg tct gaa gat gaa gtg 2068
Gln His Val Leu Ser Glu Glu Glu Leu Ser Trp Ser Glu Asp Glu Val
665 670 675 680

cag tta att gag gat gat gga caa tat gaa gtt gac gag acc tct gtg 2116
Gln Leu Ile Glu Asp Asp Gly Gln Tyr Glu Val Asp Glu Thr Ser Val
685 690 695

tcc gtt aac gtt gaa caa gat atc cag ggg tca cca cag gat gtt gtg 2164
Ser Val Asn Val Glu Gln Asp Ile Gln Gly Ser Pro Gln Asp Val Val
700 705 710

gat ccg caa gca cta aag gtg atg ctg caa gaa ctc gct gag aaa aat 2212
Asp Pro Gln Ala Leu Lys Val Met Leu Gln Glu Leu Ala Glu Lys Asn
715 720 725

tat tcg atg agg aac aag ctg ttt gtt ttt cca gag gta gtg aaa gct 2260
Tyr Ser Met Arg Asn Lys Leu Phe Val Phe Pro Glu Val Val Lys Ala
730 735 740

gat tca gtt att gat ctt tat tta aat cgt gac cta aca gct ttg gcg 2308
Asp Ser Val Ile Asp Leu Tyr Leu Asn Arg Asp Leu Thr Ala Leu Ala
745 750 755 760

aat gaa ccc gat gtc gtc atc aaa gga gca ttc aat ggt tgg aaa tgg 2356
Asn Glu Pro Asp Val Val Ile Lys Gly Ala Phe Asn Gly Trp Lys Trp
765 770 775

agg ctt ttc act gaa aga ttg cac aag agt gac ctt gga ggg gtt tgg 2404
Arg Leu Phe Thr Glu Arg Leu His Lys Ser Asp Leu Gly Gly Val Trp
780 785 790

tgg tct tgc aaa ctg tac ata ccc aag gag gcc tac aga tta gac ttt 2452
Trp Ser Cys Lys Leu Tyr Ile Pro Lys Glu Ala Tyr Arg Leu Asp Phe
795 800 805

gtg ttc ttc aac ggt cgc acg gtc tat gag aac aat ggc aac aat gat 2500
Val Phe Phe Asn Gly Arg Thr Val Tyr Glu Asn Asn Gly Asn Asn Asp

810 815 820

ttc tgt ata gga ata gaa ggc act atg aat gaa gat ctg ttt gag gat 2548
Phe Cys Ile Gly Ile Glu Gly Thr Met Asn Glu Asp Leu Phe Glu Asp
825 830 835 840

ttc ttg gtt aaa gaa aag caa agg gag ctt gag aaa ctt gcc atg gaa 2596
Phe Leu Val Lys Glu Lys Gln Arg Glu Leu Glu Lys Leu Ala Met Glu
845 850 855

gaa gct gaa agg agg aca cag act gaa gaa cag cgg cga aga aag gaa 2644
Glu Ala Glu Arg Arg Thr Gln Thr Glu Glu Gln Arg Arg Arg Lys Glu
860 865 870

gca agg gct gca gat gaa gct gtc agg gca caa gcg aag gcc gag ata 2692
Ala Arg Ala Ala Asp Glu Ala Val Arg Ala Gln Ala Lys Ala Glu Ile
875 880 885

gag atc aag aag aaa aaa ttg caa agt atg ttg agt ttg gcc aga aca 2740
Glu Ile Lys Lys Lys Lys Leu Gln Ser Met Leu Ser Leu Ala Arg Thr
890 895 900

tgt gtt gat aat ttg tgg tac ata gag gct agc aca gat aca aga gga 2788
Cys Val Asp Asn Leu Trp Tyr Ile Glu Ala Ser Thr Asp Thr Arg Gly
905 910 915 920

gat act atc agg tta tat tat aac aga aac tcg agg cca ctt gcg cat 2836
Asp Thr Ile Arg Leu Tyr Tyr Asn Arg Asn Ser Arg Pro Leu Ala His
925 930 935

agt act gag att tgg atg cat ggt ggt tac aac aat tgg aca gat gga 2884
Ser Thr Glu Ile Trp Met His Gly Gly Tyr Asn Asn Trp Thr Asp Gly
940 945 950

ctc tct att gtt gaa agc ttt gtc aag tgc aat gac aaa gac ggc gat 2932
Leu Ser Ile Val Glu Ser Phe Val Lys Cys Asn Asp Lys Asp Gly Asp
955 960 965

tgg tgg tat gca gat gtt att cca cct gaa aag gca ctt gtg ttg gac 2980
Trp Trp Tyr Ala Asp Val Ile Pro Pro Glu Lys Ala Leu Val Leu Asp
970 975 980

tgg gtt ttt gct gat ggg cca gct ggg aat gca agg aac tat gac aac 3028
Trp Val Phe Ala Asp Gly Pro Ala Gly Asn Ala Arg Asn Tyr Asp Asn
985 990 995 1000

aat gct cga caa gat ttc cat gct att ctt ccg aac aac aat gta acc 3076
Asn Ala Arg Gln Asp Phe His Ala Ile Leu Pro Asn Asn Asn Val Thr
1005 1010 1015

gag gaa ggc ttc tgg gcg caa gag gag caa aac atc tat aca agg ctt 3124
Glu Glu Gly Phe Trp Ala Gln Glu Glu Gln Asn Ile Tyr Thr Arg Leu
1020 1025 1030

ctg caa gaa agg aga gaa aag gaa gaa acc atg aaa aga aag gct gag 3172
Leu Gln Glu Arg Arg Glu Lys Glu Glu Thr Met Lys Arg Lys Ala Glu
1035 1040 1045

aga agt gca aat atc aaa gct gag atg aag gca aaa act atg cga agg 3220
Arg Ser Ala Asn Ile Lys Ala Glu Met Lys Ala Lys Thr Met Arg Arg
1050 1055 1060

ttt ctg ctt tcc cag aaa cac att gtt tat acc cga acc gnc ttg aaa 3268
Phe Leu Leu Ser Gln Lys His Ile Val Tyr Thr Arg Thr Xaa Leu Lys
1065 1070 1075 1080

```
tac gtg ccc gga acc aca gtg gat gtg cta tac aat ccc tct aac aca    3316
Tyr Val Pro Gly Thr Thr Val Asp Val Leu Tyr Asn Pro Ser Asn Thr
                1085                1090                1095

gtg cta aat gga aag tcg gag ggt tgg ttt aga tgc tcc ttt aac ctt    3364
Val Leu Asn Gly Lys Ser Glu Gly Trp Phe Arg Cys Ser Phe Asn Leu
            1100                1105                1110

tgg atg cat tca agt ggg gca ttg cca ccc cag aag atg gtg aaa tca    3412
Trp Met His Ser Ser Gly Ala Leu Pro Pro Gln Lys Met Val Lys Ser
        1115                1120                1125

ggg gat ggg ccg ctc tta aaa gca aca gtt gat gtt cca ccg gat gcc    3460
Gly Asp Gly Pro Leu Leu Lys Ala Thr Val Asp Val Pro Pro Asp Ala
    1130                1135                1140

tat atg atg gac ttt gtt ttc tcc gag tgg gaa gaa gat ggg atc tat    3508
Tyr Met Met Asp Phe Val Phe Ser Glu Trp Glu Glu Asp Gly Ile Tyr
1145                1150                1155                1160

gac aac agg aat ggg atg gac tat cat att cct gtt tct gat tca att    3556
Asp Asn Arg Asn Gly Met Asp Tyr His Ile Pro Val Ser Asp Ser Ile
                1165                1170                1175

gaa aca gag aat tac atg cgt att atc cac att gcc gtt gag atg gcc    3604
Glu Thr Glu Asn Tyr Met Arg Ile Ile His Ile Ala Val Glu Met Ala
            1180                1185                1190

ccc gtt gca aag gtt gga ggt ctt ggg gat gtt gtt aca agt ctt tca    3652
Pro Val Ala Lys Val Gly Gly Leu Gly Asp Val Val Thr Ser Leu Ser
        1195                1200                1205

cgt gcc att caa gat cta gga cat act gtc gag gtt att ctc ccg aag    3700
Arg Ala Ile Gln Asp Leu Gly His Thr Val Glu Val Ile Leu Pro Lys
    1210                1215                1220

tac gac tgt ttg aac caa agc agt gtc aag gat tta cat tta tat caa    3748
Tyr Asp Cys Leu Asn Gln Ser Ser Val Lys Asp Leu His Leu Tyr Gln
1225                1230                1235                1240

agt ttt tct tgg ggt ggt aca gaa ata aaa gta tgg gtt gga cga gtc    3796
Ser Phe Ser Trp Gly Gly Thr Glu Ile Lys Val Trp Val Gly Arg Val
                1245                1250                1255

gaa gac ctg acc gtt tac ttc ctg gaa cct caa aat ggg atg ttt ggc    3844
Glu Asp Leu Thr Val Tyr Phe Leu Glu Pro Gln Asn Gly Met Phe Gly
            1260                1265                1270

gtt gga tgt gta tat gga agg aat gat gac cgc aga ttt ggg ttc ttc    3892
Val Gly Cys Val Tyr Gly Arg Asn Asp Asp Arg Arg Phe Gly Phe Phe
        1275                1280                1285

tgt cat tct gct cta gag ttt atc ctc cag aat gaa ttt tct cca cat    3940
Cys His Ser Ala Leu Glu Phe Ile Leu Gln Asn Glu Phe Ser Pro His
    1290                1295                1300

ata ata cat tgc cat gat tgg tca agt gct ccg gtc gcc tgg cta tat    3988
Ile Ile His Cys His Asp Trp Ser Ser Ala Pro Val Ala Trp Leu Tyr
1305                1310                1315                1320

aag gaa cac tat tcc caa tcc aga atg gca agc act cgg gtt gta ttt    4036
Lys Glu His Tyr Ser Gln Ser Arg Met Ala Ser Thr Arg Val Val Phe
                1325                1330                1335
```

```
acc atc cac aat ctt gaa ttt gga gca cat tat att ggt aaa gca atg    4084
Thr Ile His Asn Leu Glu Phe Gly Ala His Tyr Ile Gly Lys Ala Met
            1340                1345                1350

aca tac tgt gat aaa gcc aca act gtt tct cct aca tat tca agg gac    4132
Thr Tyr Cys Asp Lys Ala Thr Thr Val Ser Pro Thr Tyr Ser Arg Asp
        1355                1360                1365

gtg gca ggc cat ggc gcc att gct cct cat cgt gag aaa ttc tac ggc    4180
Val Ala Gly His Gly Ala Ile Ala Pro His Arg Glu Lys Phe Tyr Gly
    1370                1375                1380

att ctc aat gga att gat cca gat atc tgg gat ccg tac act gac aat    4228
Ile Leu Asn Gly Ile Asp Pro Asp Ile Trp Asp Pro Tyr Thr Asp Asn
1385                1390                1395                1400

ttt atc ccg gtc cct tat act tgt gag aat gtt gtc gaa ggc aag aga    4276
Phe Ile Pro Val Pro Tyr Thr Cys Glu Asn Val Val Glu Gly Lys Arg
                1405                1410                1415

gct gca aaa agg gcc ttg cag cag aag ttt gga tta cag caa act gat    4324
Ala Ala Lys Arg Ala Leu Gln Gln Lys Phe Gly Leu Gln Gln Thr Asp
            1420                1425                1430

gtc cct att gtc gga atc atc acc cgt ctg aca gcc cag aag gga atc    4372
Val Pro Ile Val Gly Ile Ile Thr Arg Leu Thr Ala Gln Lys Gly Ile
        1435                1440                1445

cac ctc atc aag cac gca att cac cga act ctc gaa agc aac gga cat    4420
His Leu Ile Lys His Ala Ile His Arg Thr Leu Glu Ser Asn Gly His
    1450                1455                1460

gtg gtt ttg ctt ggt tca gct cca gat cat cga ata caa ggc gat ttt    4468
Val Val Leu Leu Gly Ser Ala Pro Asp His Arg Ile Gln Gly Asp Phe
1465                1470                1475                1480

tgc aga ttg gcc gat gct ctt cat ggt gtt tac cat ggt agg gtg aag    4516
Cys Arg Leu Ala Asp Ala Leu His Gly Val Tyr His Gly Arg Val Lys
                1485                1490                1495

ctt gtt cta acc tat gat gag cct ctt tct cac ctg ata tac gct ggc    4564
Leu Val Leu Thr Tyr Asp Glu Pro Leu Ser His Leu Ile Tyr Ala Gly
            1500                1505                1510

tcg gac ttc ata att gtt cct tca atc ttc gaa ccc tgt ggc tta aca    4612
Ser Asp Phe Ile Ile Val Pro Ser Ile Phe Glu Pro Cys Gly Leu Thr
        1515                1520                1525

caa ctt gtt gcc atg cgt tat gga tcg atc cct ata gtt cgg aaa act    4660
Gln Leu Val Ala Met Arg Tyr Gly Ser Ile Pro Ile Val Arg Lys Thr
    1530                1535                1540

gga gga ctt cac gac aca gtc ttc gac gta gac aat gat aag gac cgg    4708
Gly Gly Leu His Asp Thr Val Phe Asp Val Asp Asn Asp Lys Asp Arg
1545                1550                1555                1560

gct cgg tct ctt ggt ctt gaa cca aat ggg ttc agt ttc gac gga gcc    4756
Ala Arg Ser Leu Gly Leu Glu Pro Asn Gly Phe Ser Phe Asp Gly Ala
                1565                1570                1575

gac agc aat ggc gtg gat tat gcc ctc aac aga gca atc ggc gct tgg    4804
Asp Ser Asn Gly Val Asp Tyr Ala Leu Asn Arg Ala Ile Gly Ala Trp
            1580                1585                1590

ttc gat gcc cgt gat tgg ttc cac tcc ctg tgt aag agg gtc atg gag    4852
```

Phe Asp Ala Arg Asp Trp Phe His Ser Leu Cys Lys Arg Val Met Glu
1595 1600 1605

caa gac tgg tcg tgg aac cgg ccc gca ctg gac tac att gaa ttg tac 4900
Gln Asp Trp Ser Trp Asn Arg Pro Ala Leu Asp Tyr Ile Glu Leu Tyr
1610 1615 1620

cat gcc gct cga aaa ttc tga cacccaactg aaccaatgac aagaacaagc 4951
His Ala Ala Arg Lys Phe
1625 1630

gcattgtggg atcgactagt catacagggc tgtgcagatc gtcttgcttc agttagtgcc 5011

ctcttcagtt agttccaagc gcactacagt cgtacatagc tgaggatcct cttgcctcct 5071

accaggggga acaaagcaga aatgcatgag tgcattggga agacttttat gtatattgtt 5131

aaaaaaattt ccttttcttt tccttccctg cacctggaaa tggttaagcg catcgccgag 5191

ataagaaccg cagtgacatt ctgtgagtag ctttgtatat tctctcatct tgtgaaaact 5251

aatgttcatg ttaggctgtc tgatcatgtg gaagctttgt tatatgttac ttatggtata 5311

catcaatgat atttacattt gtggaaaaaa aaaaaaaaaa a 5352

<210> 8
<211> 1630
<212> PRT
<213> Triticum aestivum

<400> 8
Met Glu Met Ser Leu Trp Pro Arg Ser Pro Leu Cys Pro Arg Ser Arg
1 5 10 15

Gln Pro Leu Val Val Val Arg Pro Ala Gly Arg Gly Gly Leu Thr Gln
20 25 30

Pro Phe Leu Met Asn Gly Arg Phe Thr Arg Ser Arg Thr Leu Arg Cys
35 40 45

Met Val Ala Ser Ser Asp Pro Pro Asn Arg Lys Ser Arg Arg Met Val
50 55 60

Pro Pro Gln Val Lys Val Ile Ser Ser Arg Gly Tyr Thr Thr Arg Leu
65 70 75 80

Ile Val Glu Pro Ser Asn Glu Asn Thr Glu His Asn Asn Arg Asp Glu
85 90 95

Glu Thr Leu Asp Thr Tyr Asn Ala Leu Leu Ser Thr Glu Thr Ala Glu
100 105 110

Trp Thr Asp Asn Arg Glu Ala Glu Thr Ala Lys Ala Asp Ser Ser Gln
115 120 125

Asn Ala Leu Ser Ser Ser Ile Ile Gly Glu Val Asp Val Ala Asp Glu
130 135 140

Asp Ile Leu Ala Ala Asp Leu Thr Val Tyr Ser Leu Ser Ser Val Met
145 150 155 160

Lys Lys Glu Val Asp Ala Ala Asp Lys Ala Arg Val Lys Glu Asp Ala
165 170 175

```
Phe Glu Leu Asp Xaa Ala Ser Thr Thr Leu Arg Ser Val Ile Val Asp
            180                 185                 190

Val Met Asp His Xaa Trp Asp Cys Gln Glu Thr Leu Arg Ser Val Ile
        195                 200                 205

Val Asp Val Met Asp His Asn Gly Thr Val Gln Glu Thr Leu Arg Ser
    210                 215                 220

Val Ile Val Asp Val Met Asp Asp Ala Ala Asp Lys Ala Arg Val Glu
225                 230                 235                 240

Glu Asp Val Phe Glu Leu Asp Leu Ser Gly Asn Ile Ser Ser Ser Ala
                245                 250                 255

Thr Thr Val Glu Leu Asp Ala Val Asp Glu Val Gly Pro Val Gln Asp
            260                 265                 270

Lys Phe Glu Ala Thr Ser Ser Gly Asn Val Ser Asn Ser Ala Thr Val
        275                 280                 285

Arg Glu Val Asp Ala Ser Asp Glu Ala Gly Asn Asp Gln Gly Ile Phe
    290                 295                 300

Arg Ala Asp Leu Ser Gly Asn Val Phe Ser Ser Ser Thr Thr Val Glu
305                 310                 315                 320

Val Gly Ala Val Asp Glu Ala Gly Ser Ile Lys Asp Arg Phe Glu Thr
                325                 330                 335

Asp Ser Ser Gly Asn Val Ser Thr Ser Ala Pro Met Trp Asp Ala Ile
            340                 345                 350

Asp Glu Thr Val Ala Asp Gln Asp Thr Phe Glu Ala Asp Leu Ser Gly
        355                 360                 365

Asn Ala Ser Ser Cys Ala Thr Tyr Arg Glu Val Asp Asp Val Val Asp
    370                 375                 380

Glu Thr Arg Ser Glu Glu Glu Thr Phe Ala Met Asp Leu Phe Ala Ser
385                 390                 395                 400

Glu Ser Gly His Glu Lys His Met Ala Val Asp Tyr Val Gly Glu Ala
                405                 410                 415

Thr Asp Glu Glu Glu Thr Tyr Gln Gln Gln Tyr Pro Val Pro Ser Ser
            420                 425                 430

Phe Ser Met Trp Asp Lys Ala Ile Ala Lys Thr Gly Val Ser Leu Asn
        435                 440                 445

Pro Glu Leu Arg Leu Val Arg Val Glu Glu Gln Gly Lys Val Asn Phe
    450                 455                 460

Ser Asp Lys Lys Asp Leu Ser Ile Asp Asp Leu Pro Gly Gln Asn Gln
465                 470                 475                 480

Ser Ile Ile Gly Ser Tyr Lys Gln Asp Lys Ser Ile Ala Asp Val Ala
                485                 490                 495

Gly Pro Thr Gln Ser Ile Phe Gly Ser Ser Lys Gln His Arg Ser Ile
            500                 505                 510

Val Ala Phe Pro Lys Gln Asn Gln Ser Ile Val Ser Val Thr Glu Gln
        515                 520                 525
```

Lys Gln Ser Ile Val Gly Phe Arg Ser Gln Asp Leu Ser Ala Val Ser
530 535 540

Leu Pro Lys Gln Asn Val Pro Ile Val Gly Tyr Val Glu Arg Gly Ser
545 550 555 560

Asn Xaa Lys Gln Val Pro Val Val Asp Arg Gln Asp Ala Leu Tyr Val
565 570 575

Asn Gly Leu Glu Ala Lys Glu Gly Asp His Thr Ser Glu Lys Thr Asp
580 585 590

Glu Asp Ala Leu His Val Lys Phe Asn Val Asp Asn Val Leu Arg Lys
595 600 605

His Gln Ala Asp Arg Thr Gln Ala Val Glu Lys Lys Thr Trp Lys Lys
610 615 620

Val Asp Glu Glu His Leu Tyr Met Thr Glu His Gln Lys Arg Ala Ala
625 630 635 640

Glu Gly Gln Met Val Val Asn Glu Asp Glu Leu Ser Ile Thr Glu Ile
645 650 655

Gly Met Gly Arg Gly Asp Lys Ile Gln His Val Leu Ser Glu Glu Glu
660 665 670

Leu Ser Trp Ser Glu Asp Glu Val Gln Leu Ile Glu Asp Asp Gly Gln
675 680 685

Tyr Glu Val Asp Glu Thr Ser Val Ser Val Asn Val Glu Gln Asp Ile
690 695 700

Gln Gly Ser Pro Gln Asp Val Val Asp Pro Gln Ala Leu Lys Val Met
705 710 715 720

Leu Gln Glu Leu Ala Glu Lys Asn Tyr Ser Met Arg Asn Lys Leu Phe
725 730 735

Val Phe Pro Glu Val Val Lys Ala Asp Ser Val Ile Asp Leu Tyr Leu
740 745 750

Asn Arg Asp Leu Thr Ala Leu Ala Asn Glu Pro Asp Val Val Ile Lys
755 760 765

Gly Ala Phe Asn Gly Trp Lys Trp Arg Leu Phe Thr Glu Arg Leu His
770 775 780

Lys Ser Asp Leu Gly Gly Val Trp Trp Ser Cys Lys Leu Tyr Ile Pro
785 790 795 800

Lys Glu Ala Tyr Arg Leu Asp Phe Val Phe Phe Asn Gly Arg Thr Val
805 810 815

Tyr Glu Asn Asn Gly Asn Asn Asp Phe Cys Ile Gly Ile Glu Gly Thr
820 825 830

Met Asn Glu Asp Leu Phe Glu Asp Phe Leu Val Lys Glu Lys Gln Arg
835 840 845

Glu Leu Glu Lys Leu Ala Met Glu Glu Ala Glu Arg Arg Thr Gln Thr
850 855 860

Glu Glu Gln Arg Arg Arg Lys Glu Ala Arg Ala Ala Asp Glu Ala Val

865 870 875 880

Arg Ala Gln Ala Lys Ala Glu Ile Glu Ile Lys Lys Lys Lys Leu Gln
885 890 895

Ser Met Leu Ser Leu Ala Arg Thr Cys Val Asp Asn Leu Trp Tyr Ile
900 905 910

Glu Ala Ser Thr Asp Thr Arg Gly Asp Thr Ile Arg Leu Tyr Tyr Asn
915 920 925

Arg Asn Ser Arg Pro Leu Ala His Ser Thr Glu Ile Trp Met His Gly
930 935 940

Gly Tyr Asn Asn Trp Thr Asp Gly Leu Ser Ile Val Glu Ser Phe Val
945 950 955 960

Lys Cys Asn Asp Lys Asp Gly Asp Trp Trp Tyr Ala Asp Val Ile Pro
965 970 975

Pro Glu Lys Ala Leu Val Leu Asp Trp Val Phe Ala Asp Gly Pro Ala
980 985 990

Gly Asn Ala Arg Asn Tyr Asp Asn Asn Ala Arg Gln Asp Phe His Ala
995 1000 1005

Ile Leu Pro Asn Asn Asn Val Thr Glu Glu Gly Phe Trp Ala Gln Glu
1010 1015 1020

Glu Gln Asn Ile Tyr Thr Arg Leu Leu Gln Glu Arg Arg Glu Lys Glu
025 1030 1035 1040

Glu Thr Met Lys Arg Lys Ala Glu Arg Ser Ala Asn Ile Lys Ala Glu
1045 1050 1055

Met Lys Ala Lys Thr Met Arg Arg Phe Leu Leu Ser Gln Lys His Ile
1060 1065 1070

Val Tyr Thr Arg Thr Xaa Leu Lys Tyr Val Pro Gly Thr Thr Val Asp
1075 1080 1085

Val Leu Tyr Asn Pro Ser Asn Thr Val Leu Asn Gly Lys Ser Glu Gly
1090 1095 1100

Trp Phe Arg Cys Ser Phe Asn Leu Trp Met His Ser Ser Gly Ala Leu
105 1110 1115 1120

Pro Pro Gln Lys Met Val Lys Ser Gly Asp Gly Pro Leu Leu Lys Ala
1125 1130 1135

Thr Val Asp Val Pro Pro Asp Ala Tyr Met Met Asp Phe Val Phe Ser
1140 1145 1150

Glu Trp Glu Glu Asp Gly Ile Tyr Asp Asn Arg Asn Gly Met Asp Tyr
1155 1160 1165

His Ile Pro Val Ser Asp Ser Ile Glu Thr Glu Asn Tyr Met Arg Ile
1170 1175 1180

Ile His Ile Ala Val Glu Met Ala Pro Val Ala Lys Val Gly Gly Leu
185 1190 1195 1200

Gly Asp Val Val Thr Ser Leu Ser Arg Ala Ile Gln Asp Leu Gly His
1205 1210 1215

Thr Val Glu Val Ile Leu Pro Lys Tyr Asp Cys Leu Asn Gln Ser Ser
1220 1225 1230

Val Lys Asp Leu His Leu Tyr Gln Ser Phe Ser Trp Gly Gly Thr Glu
1235 1240 1245

Ile Lys Val Trp Val Gly Arg Val Glu Asp Leu Thr Val Tyr Phe Leu
1250 1255 1260

Glu Pro Gln Asn Gly Met Phe Gly Val Gly Cys Val Tyr Gly Arg Asn
265 1270 1275 1280

Asp Asp Arg Arg Phe Gly Phe Phe Cys His Ser Ala Leu Glu Phe Ile
1285 1290 1295

Leu Gln Asn Glu Phe Ser Pro His Ile Ile His Cys His Asp Trp Ser
1300 1305 1310

Ser Ala Pro Val Ala Trp Leu Tyr Lys Glu His Tyr Ser Gln Ser Arg
1315 1320 1325

Met Ala Ser Thr Arg Val Val Phe Thr Ile His Asn Leu Glu Phe Gly
1330 1335 1340

Ala His Tyr Ile Gly Lys Ala Met Thr Tyr Cys Asp Lys Ala Thr Thr
345 1350 1355 1360

Val Ser Pro Thr Tyr Ser Arg Asp Val Ala Gly His Gly Ala Ile Ala
1365 1370 1375

Pro His Arg Glu Lys Phe Tyr Gly Ile Leu Asn Gly Ile Asp Pro Asp
1380 1385 1390

Ile Trp Asp Pro Tyr Thr Asp Asn Phe Ile Pro Val Pro Tyr Thr Cys
1395 1400 1405

Glu Asn Val Val Glu Gly Lys Arg Ala Ala Lys Arg Ala Leu Gln Gln
1410 1415 1420

Lys Phe Gly Leu Gln Gln Thr Asp Val Pro Ile Val Gly Ile Ile Thr
425 1430 1435 1440

Arg Leu Thr Ala Gln Lys Gly Ile His Leu Ile Lys His Ala Ile His
1445 1450 1455

Arg Thr Leu Glu Ser Asn Gly His Val Val Leu Leu Gly Ser Ala Pro
1460 1465 1470

Asp His Arg Ile Gln Gly Asp Phe Cys Arg Leu Ala Asp Ala Leu His
1475 1480 1485

Gly Val Tyr His Gly Arg Val Lys Leu Val Leu Thr Tyr Asp Glu Pro
1490 1495 1500

Leu Ser His Leu Ile Tyr Ala Gly Ser Asp Phe Ile Ile Val Pro Ser
505 1510 1515 1520

Ile Phe Glu Pro Cys Gly Leu Thr Gln Leu Val Ala Met Arg Tyr Gly
1525 1530 1535

Ser Ile Pro Ile Val Arg Lys Thr Gly Gly Leu His Asp Thr Val Phe
1540 1545 1550

Asp Val Asp Asn Asp Lys Asp Arg Ala Arg Ser Leu Gly Leu Glu Pro
1555 1560 1565

Asn Gly Phe Ser Phe Asp Gly Ala Asp Ser Asn Gly Val Asp Tyr Ala
1570 1575 1580

Leu Asn Arg Ala Ile Gly Ala Trp Phe Asp Ala Arg Asp Trp Phe His
585 1590 1595 1600

Ser Leu Cys Lys Arg Val Met Glu Gln Asp Trp Ser Trp Asn Arg Pro
1605 1610 1615

Ala Leu Asp Tyr Ile Glu Leu Tyr His Ala Ala Arg Lys Phe
1620 1625 1630

<210> 9
<211> 3621
<212> DNA
<213> Triticum aestivum

<220>
<221> CDS
<222> (1)..(3180)

<400> 9
gat gca ttg tat gtg aat gga ctg gaa gct aag gag gga gat cac aca 48
Asp Ala Leu Tyr Val Asn Gly Leu Glu Ala Lys Glu Gly Asp His Thr
1 5 10 15

tcc gag aaa act gat gag gat gcg ctt cat gta aag ttt aat gtt gac 96
Ser Glu Lys Thr Asp Glu Asp Ala Leu His Val Lys Phe Asn Val Asp
20 25 30

aat gtg ttg cgg aag cat cag gca gat aga acc caa gca gtg gaa aag 144
Asn Val Leu Arg Lys His Gln Ala Asp Arg Thr Gln Ala Val Glu Lys
35 40 45

aaa act tgg aag aaa gtt gat gag gaa cat ctt tac atg act gaa cat 192
Lys Thr Trp Lys Lys Val Asp Glu Glu His Leu Tyr Met Thr Glu His
50 55 60

cag aaa cgt gct gcc gaa gga cag atg gta gtt aac gag gat gag ctt 240
Gln Lys Arg Ala Ala Glu Gly Gln Met Val Val Asn Glu Asp Glu Leu
65 70 75 80

tct ata act gaa att gga atg ggg aga ggt gat aaa att cag cat gtg 288
Ser Ile Thr Glu Ile Gly Met Gly Arg Gly Asp Lys Ile Gln His Val
85 90 95

ctt tct gag gaa gag ctt tca tgg tct gaa gat gaa gtg cag tta att 336
Leu Ser Glu Glu Glu Leu Ser Trp Ser Glu Asp Glu Val Gln Leu Ile
100 105 110

gag gat gat gga caa tat gaa gtt gac gag acc tct gtg tcc gtt aac 384
Glu Asp Asp Gly Gln Tyr Glu Val Asp Glu Thr Ser Val Ser Val Asn
115 120 125

gtt gaa caa gat atc cag ggg tca cca cag gat gtt gtg gat ccg caa 432
Val Glu Gln Asp Ile Gln Gly Ser Pro Gln Asp Val Val Asp Pro Gln
130 135 140

gca cta aag gtg atg ctg caa gaa ctc gct gag aaa aat tat tcg atg 480
Ala Leu Lys Val Met Leu Gln Glu Leu Ala Glu Lys Asn Tyr Ser Met
145 150 155 160

agg aac aag ctg ttt gtt ttt cca gag gta gtg aaa gct gat tca gtt 528

```
Arg Asn Lys Leu Phe Val Phe Pro Glu Val Val Lys Ala Asp Ser Val
                165                 170                 175

att gat ctt tat tta aat cgt gac cta aca gct ttg gcg aat gaa ccc     576
Ile Asp Leu Tyr Leu Asn Arg Asp Leu Thr Ala Leu Ala Asn Glu Pro
            180                 185                 190

gat gtc gtc atc aaa gga gca ttc aat ggt tgg aaa tgg agg ctt ttc     624
Asp Val Val Ile Lys Gly Ala Phe Asn Gly Trp Lys Trp Arg Leu Phe
        195                 200                 205

act gaa aga ttg cac aag agt gac ctt gga ggg gtt tgg tgg tct tgc     672
Thr Glu Arg Leu His Lys Ser Asp Leu Gly Gly Val Trp Trp Ser Cys
    210                 215                 220

aaa ctg tac ata ccc aag gag gcc tac aga tta gac ttt gtg ttc ttc     720
Lys Leu Tyr Ile Pro Lys Glu Ala Tyr Arg Leu Asp Phe Val Phe Phe
225                 230                 235                 240

aac ggt cgc acg gtc tat gag aac aat ggc aac aat gat ttc tgt ata     768
Asn Gly Arg Thr Val Tyr Glu Asn Asn Gly Asn Asn Asp Phe Cys Ile
                245                 250                 255

gga ata gaa ggc act atg aat gaa gat ctg ttt gag gat ttc ttg gtt     816
Gly Ile Glu Gly Thr Met Asn Glu Asp Leu Phe Glu Asp Phe Leu Val
            260                 265                 270

aaa gaa aag caa agg gag ctt gag aaa ctt gcc atg gaa gaa gct gaa     864
Lys Glu Lys Gln Arg Glu Leu Glu Lys Leu Ala Met Glu Glu Ala Glu
        275                 280                 285

agg agg aca cag act gaa gaa cag cgg cga aga aag gaa gca agg gct     912
Arg Arg Thr Gln Thr Glu Glu Gln Arg Arg Arg Lys Glu Ala Arg Ala
    290                 295                 300

gca gat gaa gct gtc agg gca caa gcg aag gcc gag ata gag atc aag     960
Ala Asp Glu Ala Val Arg Ala Gln Ala Lys Ala Glu Ile Glu Ile Lys
305                 310                 315                 320

aag aaa aaa ttg caa agt atg ttg agt ttg gcc aga aca tgt gtt gat     1008
Lys Lys Lys Leu Gln Ser Met Leu Ser Leu Ala Arg Thr Cys Val Asp
                325                 330                 335

aat ttg tgg tac ata gag gct agc aca gat aca aga gga gat act atc     1056
Asn Leu Trp Tyr Ile Glu Ala Ser Thr Asp Thr Arg Gly Asp Thr Ile
            340                 345                 350

agg tta tat tat aac aga aac tcg agg cca ctt gcg cat agt act gag     1104
Arg Leu Tyr Tyr Asn Arg Asn Ser Arg Pro Leu Ala His Ser Thr Glu
        355                 360                 365

att tgg atg cat ggt ggt tac aac aat tgg tca gat gga ctc tct att     1152
Ile Trp Met His Gly Gly Tyr Asn Asn Trp Ser Asp Gly Leu Ser Ile
    370                 375                 380

gtt gaa agc ttt gtc aag tgc aat gac aaa gac ggc gat tgg tgg tat     1200
Val Glu Ser Phe Val Lys Cys Asn Asp Lys Asp Gly Asp Trp Trp Tyr
385                 390                 395                 400

gca gat gtt att cca cct gaa aag gca ctt gtg ttg gac tgg gtt ttt     1248
Ala Asp Val Ile Pro Pro Glu Lys Ala Leu Val Leu Asp Trp Val Phe
                405                 410                 415

gct gat ggg cca gct ggg aat gca agg aac tat gac aac aat gct cga     1296
Ala Asp Gly Pro Ala Gly Asn Ala Arg Asn Tyr Asp Asn Asn Ala Arg
```

```
            420                   425                   430

caa gat ttc cat gct att ctt ccg aac aac aat gta acc gag gaa ggc    1344
Gln Asp Phe His Ala Ile Leu Pro Asn Asn Asn Val Thr Glu Glu Gly
        435                 440                 445

ttc tgg gcg caa gag gag caa aac atc tat aca agg ctt ctg caa gaa    1392
Phe Trp Ala Gln Glu Glu Gln Asn Ile Tyr Thr Arg Leu Leu Gln Glu
    450                 455                 460

agg aga gaa aag gaa gaa acc atg aaa aga aag gct gag aga agt gca    1440
Arg Arg Glu Lys Glu Glu Thr Met Lys Arg Lys Ala Glu Arg Ser Ala
465                 470                 475                 480

aat atc aaa gct gag atg aag gca aaa act atg cga agg ttt ctg ctt    1488
Asn Ile Lys Ala Glu Met Lys Ala Lys Thr Met Arg Arg Phe Leu Leu
                485                 490                 495

tcc cag aaa cac att gtt tat acc cga acc gnc ttg aaa tac gtg ccc    1536
Ser Gln Lys His Ile Val Tyr Thr Arg Thr Xaa Leu Lys Tyr Val Pro
            500                 505                 510

gga acc aca gtg gat gtg cta tac aat ccc tct aac aca gtg cta aat    1584
Gly Thr Thr Val Asp Val Leu Tyr Asn Pro Ser Asn Thr Val Leu Asn
        515                 520                 525

gga aag tcg gag ggt tgg ttt aga tgc tcc ttt aac ctt tgg atg cat    1632
Gly Lys Ser Glu Gly Trp Phe Arg Cys Ser Phe Asn Leu Trp Met His
    530                 535                 540

tca agt ggg gca ttg cca ccc cag aag atg gtg aaa tca ggg gat ggg    1680
Ser Ser Gly Ala Leu Pro Pro Gln Lys Met Val Lys Ser Gly Asp Gly
545                 550                 555                 560

ccg ctc tta aaa gca aca gtt gat gtt cca ccg gat gcc tat atg atg    1728
Pro Leu Leu Lys Ala Thr Val Asp Val Pro Pro Asp Ala Tyr Met Met
                565                 570                 575

gac ttt gtt ttc tcc gag tgg gaa gaa gat ggg atc tat gac aac agg    1776
Asp Phe Val Phe Ser Glu Trp Glu Glu Asp Gly Ile Tyr Asp Asn Arg
            580                 585                 590

aat ggg atg gac tat cat att cct gtt tct gat tca att gaa aca gag    1824
Asn Gly Met Asp Tyr His Ile Pro Val Ser Asp Ser Ile Glu Thr Glu
        595                 600                 605

aat tac atg cgt att atc cac att gcc gtt gag atg gcc ccc gtt gca    1872
Asn Tyr Met Arg Ile Ile His Ile Ala Val Glu Met Ala Pro Val Ala
    610                 615                 620

aag gtt gga ggt ctt ggg gat gtt gtt aca agt ctt tca cgt gcc att    1920
Lys Val Gly Gly Leu Gly Asp Val Val Thr Ser Leu Ser Arg Ala Ile
625                 630                 635                 640

caa gat cta gga cat act gtc gag gtt att ctc ccg aag tac gac tgt    1968
Gln Asp Leu Gly His Thr Val Glu Val Ile Leu Pro Lys Tyr Asp Cys
                645                 650                 655

ttg aac caa agc agt gtc aag gat tta cat tta tat caa agt ttt tct    2016
Leu Asn Gln Ser Ser Val Lys Asp Leu His Leu Tyr Gln Ser Phe Ser
            660                 665                 670

tgg ggt ggt aca gaa ata aaa gta tgg gtt gga cga gtc gaa gac ctg    2064
Trp Gly Gly Thr Glu Ile Lys Val Trp Val Gly Arg Val Glu Asp Leu
        675                 680                 685
```

```
acc gtt tac ttc ctg gaa cct caa aat ggg atg ttt ggc gtt gga tgt     2112
Thr Val Tyr Phe Leu Glu Pro Gln Asn Gly Met Phe Gly Val Gly Cys
    690                 695                 700

gta tat gga agg aat gat gac cgc aga ttt ggg ttc ttc tgt cat tct     2160
Val Tyr Gly Arg Asn Asp Asp Arg Arg Phe Gly Phe Phe Cys His Ser
705                 710                 715                 720

gct cta gag ttt atc ctc cag aat gaa ttt tct cca cat ata ata cat     2208
Ala Leu Glu Phe Ile Leu Gln Asn Glu Phe Ser Pro His Ile Ile His
                725                 730                 735

tgc cat gat tgg tca agt gct ccg gtc gcc tgg cta tat aag gaa cac     2256
Cys His Asp Trp Ser Ser Ala Pro Val Ala Trp Leu Tyr Lys Glu His
            740                 745                 750

tat tcc caa tcc aga atg gca agc act cgg gtt gta ttt acc atc cac     2304
Tyr Ser Gln Ser Arg Met Ala Ser Thr Arg Val Val Phe Thr Ile His
        755                 760                 765

aat ctt gaa ttt gga gca cat tat att ggt aaa gca atg aca tac tgt     2352
Asn Leu Glu Phe Gly Ala His Tyr Ile Gly Lys Ala Met Thr Tyr Cys
    770                 775                 780

gat aaa gcc aca act gtt tct cct aca tat tca agg gac gtg gca ggc     2400
Asp Lys Ala Thr Thr Val Ser Pro Thr Tyr Ser Arg Asp Val Ala Gly
785                 790                 795                 800

cat ggc gcc att gct cct cat cgt gag aaa ttc tac ggc att ctc aat     2448
His Gly Ala Ile Ala Pro His Arg Glu Lys Phe Tyr Gly Ile Leu Asn
                805                 810                 815

gga att gat cca gat atc tgg gat ccg tac act gac aat ttt atc ccg     2496
Gly Ile Asp Pro Asp Ile Trp Asp Pro Tyr Thr Asp Asn Phe Ile Pro
            820                 825                 830

gtc cct tat act tgt gag aat gtt gtc gaa ggc aag agg gct gca aaa     2544
Val Pro Tyr Thr Cys Glu Asn Val Val Glu Gly Lys Arg Ala Ala Lys
        835                 840                 845

agg gcc ttg cag cag aag ttt gga tta cag caa act gat gtc cct att     2592
Arg Ala Leu Gln Gln Lys Phe Gly Leu Gln Gln Thr Asp Val Pro Ile
    850                 855                 860

gtc gga atc atc acc cgt ctg aca gca cag aag gga atc cac ctc atc     2640
Val Gly Ile Ile Thr Arg Leu Thr Ala Gln Lys Gly Ile His Leu Ile
865                 870                 875                 880

aag cac gca att cac cga acc ctc gag agc aat gga caa gtg gtt ttg     2688
Lys His Ala Ile His Arg Thr Leu Glu Ser Asn Gly Gln Val Val Leu
                885                 890                 895

ctt ggt tca gct cca gat cat cga ata caa ggc gat ttt tgc aga ttg     2736
Leu Gly Ser Ala Pro Asp His Arg Ile Gln Gly Asp Phe Cys Arg Leu
            900                 905                 910

gcc gat gct ctt cac ggt gtt tac cat ggt agg gtg aag ctt gtt cta     2784
Ala Asp Ala Leu His Gly Val Tyr His Gly Arg Val Lys Leu Val Leu
        915                 920                 925

acc tac gat gag cct ctt tct cac ctg ata tac gct ggc tcc gac ttc     2832
Thr Tyr Asp Glu Pro Leu Ser His Leu Ile Tyr Ala Gly Ser Asp Phe
    930                 935                 940
```

```
att att gtc cct tca atc ttt gaa ccc tgt ggc tta aca caa ctt gtt    2880
Ile Ile Val Pro Ser Ile Phe Glu Pro Cys Gly Leu Thr Gln Leu Val
945                 950                 955                 960

gcc atg cgt tat gga tcg atc cct ata gtt cgg aaa acc gga gga ctt    2928
Ala Met Arg Tyr Gly Ser Ile Pro Ile Val Arg Lys Thr Gly Gly Leu
                965                 970                 975

tac gac act gtc ttc gac gta gac aat gat aag gac cgg gct cgg tct    2976
Tyr Asp Thr Val Phe Asp Val Asp Asn Asp Lys Asp Arg Ala Arg Ser
            980                 985                 990

ctt ggt ctt gaa cca aat ggg ttc agt ttc gac gga gcc gac agc aat    3024
Leu Gly Leu Glu Pro Asn Gly Phe Ser Phe Asp Gly Ala Asp Ser Asn
        995                 1000                1005

ggc gtg gat tat gcc ctc aac aga gca atc ggc gct tgg ttc gat gcc    3072
Gly Val Asp Tyr Ala Leu Asn Arg Ala Ile Gly Ala Trp Phe Asp Ala
    1010                1015                1020

cgt gat tgg ttc cac tcc ctg tgt aag agg gtc atg gag caa gac tgg    3120
Arg Asp Trp Phe His Ser Leu Cys Lys Arg Val Met Glu Gln Asp Trp
1025                1030                1035                1040

tcg tgg aac cgg cct gca ctg gac tac att gaa ttg tac cat gcc gct    3168
Ser Trp Asn Arg Pro Ala Leu Asp Tyr Ile Glu Leu Tyr His Ala Ala
                1045                1050                1055

cga aaa ttc tga cacccaactg aaccaatggc aagaacaagc gcattgtggg        3220
Arg Lys Phe
            1060

atcgactaca gtcatacagg gctgtgcaga tcgtcttgct tcagttagtg ccctcttcag 3280

ttagttccaa gcgcactaca gtcgtacata gctgaggatc ctcttgcctc ctccaccagg 3340

ggaaacaaag cagaaatgca taagtgcatt gggaagactt ttatgtatat tgttaaattt 3400

ttccttttct tttccttccc tgcacctgga aatggttaag cgcatcgccg agataagaac 3460

cacagtaaca ttctgtgagt agctttgtat attctctcat cttgtgaaaa ctaatgtgca 3520

tgttaggctc tctgatcatg tggaagcttt gttatatgtt acttatggtt atatggtata 3580

catcaatgat atttacattt gtggaaaaaa aaaaaaaaaa a                      3621
```

```
<210> 10
<211> 1059
<212> PRT
<213> Triticum aestivum

<400> 10
Asp Ala Leu Tyr Val Asn Gly Leu Glu Ala Lys Glu Gly Asp His Thr
 1               5                  10                  15

Ser Glu Lys Thr Asp Glu Asp Ala Leu His Val Lys Phe Asn Val Asp
            20                  25                  30

Asn Val Leu Arg Lys His Gln Ala Asp Arg Thr Gln Ala Val Glu Lys
        35                  40                  45

Lys Thr Trp Lys Lys Val Asp Glu Glu His Leu Tyr Met Thr Glu His
    50                  55                  60
```

```
Gln Lys Arg Ala Ala Glu Gly Gln Met Val Val Asn Glu Asp Glu Leu
65                  70                  75                  80

Ser Ile Thr Glu Ile Gly Met Gly Arg Gly Asp Lys Ile Gln His Val
                85                  90                  95

Leu Ser Glu Glu Glu Leu Ser Trp Ser Glu Asp Glu Val Gln Leu Ile
            100                 105                 110

Glu Asp Asp Gly Gln Tyr Glu Val Asp Glu Thr Ser Val Ser Val Asn
        115                 120                 125

Val Glu Gln Asp Ile Gln Gly Ser Pro Gln Asp Val Val Asp Pro Gln
    130                 135                 140

Ala Leu Lys Val Met Leu Gln Glu Leu Ala Glu Lys Asn Tyr Ser Met
145                 150                 155                 160

Arg Asn Lys Leu Phe Val Phe Pro Glu Val Val Lys Ala Asp Ser Val
                165                 170                 175

Ile Asp Leu Tyr Leu Asn Arg Asp Leu Thr Ala Leu Ala Asn Glu Pro
            180                 185                 190

Asp Val Val Ile Lys Gly Ala Phe Asn Gly Trp Lys Trp Arg Leu Phe
        195                 200                 205

Thr Glu Arg Leu His Lys Ser Asp Leu Gly Gly Val Trp Trp Ser Cys
    210                 215                 220

Lys Leu Tyr Ile Pro Lys Glu Ala Tyr Arg Leu Asp Phe Val Phe Phe
225                 230                 235                 240

Asn Gly Arg Thr Val Tyr Glu Asn Asn Gly Asn Asn Asp Phe Cys Ile
                245                 250                 255

Gly Ile Glu Gly Thr Met Asn Glu Asp Leu Phe Glu Asp Phe Leu Val
            260                 265                 270

Lys Glu Lys Gln Arg Glu Leu Glu Lys Leu Ala Met Glu Glu Ala Glu
        275                 280                 285

Arg Arg Thr Gln Thr Glu Glu Gln Arg Arg Arg Lys Glu Ala Arg Ala
    290                 295                 300

Ala Asp Glu Ala Val Arg Ala Gln Ala Lys Ala Glu Ile Glu Ile Lys
305                 310                 315                 320

Lys Lys Lys Leu Gln Ser Met Leu Ser Leu Ala Arg Thr Cys Val Asp
                325                 330                 335

Asn Leu Trp Tyr Ile Glu Ala Ser Thr Asp Thr Arg Gly Asp Thr Ile
            340                 345                 350

Arg Leu Tyr Tyr Asn Arg Asn Ser Arg Pro Leu Ala His Ser Thr Glu
        355                 360                 365

Ile Trp Met His Gly Gly Tyr Asn Asn Trp Ser Asp Gly Leu Ser Ile
    370                 375                 380

Val Glu Ser Phe Val Lys Cys Asn Asp Lys Asp Gly Asp Trp Trp Tyr
385                 390                 395                 400

Ala Asp Val Ile Pro Pro Glu Lys Ala Leu Val Leu Asp Trp Val Phe
                405                 410                 415
```

```
Ala Asp Gly Pro Ala Gly Asn Ala Arg Asn Tyr Asp Asn Asn Ala Arg
            420                 425                 430

Gln Asp Phe His Ala Ile Leu Pro Asn Asn Asn Val Thr Glu Glu Gly
        435                 440                 445

Phe Trp Ala Gln Glu Glu Gln Asn Ile Tyr Thr Arg Leu Leu Gln Glu
    450                 455                 460

Arg Arg Glu Lys Glu Glu Thr Met Lys Arg Lys Ala Glu Arg Ser Ala
465                 470                 475                 480

Asn Ile Lys Ala Glu Met Lys Ala Lys Thr Met Arg Arg Phe Leu Leu
                485                 490                 495

Ser Gln Lys His Ile Val Tyr Thr Arg Thr Xaa Leu Lys Tyr Val Pro
            500                 505                 510

Gly Thr Thr Val Asp Val Leu Tyr Asn Pro Ser Asn Thr Val Leu Asn
        515                 520                 525

Gly Lys Ser Glu Gly Trp Phe Arg Cys Ser Phe Asn Leu Trp Met His
    530                 535                 540

Ser Ser Gly Ala Leu Pro Pro Gln Lys Met Val Lys Ser Gly Asp Gly
545                 550                 555                 560

Pro Leu Leu Lys Ala Thr Val Asp Val Pro Pro Asp Ala Tyr Met Met
                565                 570                 575

Asp Phe Val Phe Ser Glu Trp Glu Glu Asp Gly Ile Tyr Asp Asn Arg
            580                 585                 590

Asn Gly Met Asp Tyr His Ile Pro Val Ser Asp Ser Ile Glu Thr Glu
        595                 600                 605

Asn Tyr Met Arg Ile Ile His Ile Ala Val Glu Met Ala Pro Val Ala
    610                 615                 620

Lys Val Gly Gly Leu Gly Asp Val Val Thr Ser Leu Ser Arg Ala Ile
625                 630                 635                 640

Gln Asp Leu Gly His Thr Val Glu Val Ile Leu Pro Lys Tyr Asp Cys
                645                 650                 655

Leu Asn Gln Ser Ser Val Lys Asp Leu His Leu Tyr Gln Ser Phe Ser
            660                 665                 670

Trp Gly Gly Thr Glu Ile Lys Val Trp Val Gly Arg Val Glu Asp Leu
        675                 680                 685

Thr Val Tyr Phe Leu Glu Pro Gln Asn Gly Met Phe Gly Val Gly Cys
    690                 695                 700

Val Tyr Gly Arg Asn Asp Asp Arg Arg Phe Gly Phe Phe Cys His Ser
705                 710                 715                 720

Ala Leu Glu Phe Ile Leu Gln Asn Glu Phe Ser Pro His Ile Ile His
                725                 730                 735

Cys His Asp Trp Ser Ser Ala Pro Val Ala Trp Leu Tyr Lys Glu His
            740                 745                 750

Tyr Ser Gln Ser Arg Met Ala Ser Thr Arg Val Val Phe Thr Ile His
```

755 760 765

Asn Leu Glu Phe Gly Ala His Tyr Ile Gly Lys Ala Met Thr Tyr Cys
770 775 780

Asp Lys Ala Thr Thr Val Ser Pro Thr Tyr Ser Arg Asp Val Ala Gly
785 790 795 800

His Gly Ala Ile Ala Pro His Arg Glu Lys Phe Tyr Gly Ile Leu Asn
805 810 815

Gly Ile Asp Pro Asp Ile Trp Asp Pro Tyr Thr Asp Asn Phe Ile Pro
820 825 830

Val Pro Tyr Thr Cys Glu Asn Val Val Glu Gly Lys Arg Ala Ala Lys
835 840 845

Arg Ala Leu Gln Gln Lys Phe Gly Leu Gln Gln Thr Asp Val Pro Ile
850 855 860

Val Gly Ile Ile Thr Arg Leu Thr Ala Gln Lys Gly Ile His Leu Ile
865 870 875 880

Lys His Ala Ile His Arg Thr Leu Glu Ser Asn Gly Gln Val Val Leu
885 890 895

Leu Gly Ser Ala Pro Asp His Arg Ile Gln Gly Asp Phe Cys Arg Leu
900 905 910

Ala Asp Ala Leu His Gly Val Tyr His Gly Arg Val Lys Leu Val Leu
915 920 925

Thr Tyr Asp Glu Pro Leu Ser His Leu Ile Tyr Ala Gly Ser Asp Phe
930 935 940

Ile Ile Val Pro Ser Ile Phe Glu Pro Cys Gly Leu Thr Gln Leu Val
945 950 955 960

Ala Met Arg Tyr Gly Ser Ile Pro Ile Val Arg Lys Thr Gly Gly Leu
965 970 975

Tyr Asp Thr Val Phe Asp Val Asp Asn Asp Lys Asp Arg Ala Arg Ser
980 985 990

Leu Gly Leu Glu Pro Asn Gly Phe Ser Phe Asp Gly Ala Asp Ser Asn
995 1000 1005

Gly Val Asp Tyr Ala Leu Asn Arg Ala Ile Gly Ala Trp Phe Asp Ala
1010 1015 1020

Arg Asp Trp Phe His Ser Leu Cys Lys Arg Val Met Glu Gln Asp Trp
025 1030 1035 1040

Ser Trp Asn Arg Pro Ala Leu Asp Tyr Ile Glu Leu Tyr His Ala Ala
1045 1050 1055

Arg Lys Phe

<210> 11
<211> 728
<212> DNA
<213> Triticum sp.

<400> 11

gatcttgaac ggcacgtgaa agacttgtaa caacatcccc gagacctcca acctatgaga 60

tcatcgatca tgacagagca tagtattatg gcatagaatg aaaaaaaggc ataaggtgat 120

gagatctcca cgccagagcg ttgtattcca attttagttc tttccccgtg aggaggggag 180

gctaggcggg cgaggcagag gggatagggc agtcgccgct gcgtggtgga ctgactggtg 240

tggtgggtgg tgggttttgc gggcggggtt tagtaggttc ccggaaatgg agatggctct 300

ccggccacgg agccctctgt gccctcggag cagtcagccg ctcgtcgtcg tccggccggc 360

cggccgcggc ggcggcctcg cgcaggtacg ggtgattatg gttcttgatt cggtcggttc 420

acggaatgtt gtttgatttg gttctgtccc gggtcaggtt catagtgatt ttattccgca 480

aaaaaaaaag gtttatagtg attttgattt ctttcatctc gggaacattt ttatatctgg 540

gagtcaaagg gcattggttt tgatttgcat gcggaacata ttggttattt attaatgtgg 600

tgagctggaa ttcatactgc ttaaaacgac gtgattttaa ttgctggaag aggtaaagaa 660

catgaattct tgttatattt gttaaaaaaa atcccctgtt ctagcgtttc aatctgcatg 720

atcatgga 728

<210> 12
<211> 2446
<212> DNA
<213> Triticum sp.

<400> 12
gtgggtctat aaaagacagg tttgagcgga ttcgtcagga aatgtttcaa caagtgcgac 60

gatgtgggat gcaattgatg aaaccgtggc ttgatcaaga cgcagttgag gcggatttgt 120

cgggaaatgc ttcaagctgc gcgacataca gagaagtgga tgatgtggtg gatgaaacta 180

gatcagaaga ggaaacattt gcgatggatt tgtttgcaag tgaatcaggc catgagaaac 240

atatggcagt ggatcatgtg ggtgaagcta ccgatgaaga agagacttac caacagcaat 300

atccagtacc gtcttcattc tctatgtggg acaaggctat tgctaaaaca ggtgtaagtt 360

tgaatcctga gctgcgactt gtcagggttg aagaacaagg caaagtaaat tttagtgata 420

aaaaagacct gtcaattgat gatttaccag gacaaaacca atcgatcatt ggttcctata 480

aacaagataa atcaattgct gatgttgcgg gaccgaccca atcaattttt ggttctagta 540

aacaacaccg gtcaattgtt gctttcccca aacaaaacca gtcaattgtt agtgtcactg 600

agcaaaagca gtccatagtt ggattccgta gtcaagatct ttcggctgtt agtctcccta 660

aacaaaacgt accaattgtt ggtacgtcga gagagggtca aacaaagcaa gttcctgttg 720

ttgatagaca ggatgcgttg tatgtgaatg gactggaagc taaggaggga gatcacacat 780

ccgagaaaac cgatgaggat gtgcttcatg taaaatttaa tgttgacaat gtgttgcgga 840

agcatcaggc agatagaacc caagcagtgg aaacgataac ttggaagaaa gttgatgagg 900

aacatcttta catgactgaa catcagatag gtgctgccga aggacagatg gtagttaacg 960

aggatgagct ttctataact gaaattggaa tggggagagg tgataaaatt cagcatgtgc 1020

tttctgagga agagctttca tggtctgaag atgaagtgca gttaattgag gatgatggac 1080

aatatgaagt tgatgagacc tctgtgtccg ttaacgttga acaagatatc caggggtcac 1140

cacaggatgt tgtggatccg caagcactaa aggtgatgct gcaagaactc gctgagaaaa 1200

attattcgat gaggaacaag ctgtttgttt ttccagaggt agtgaaagct gattcagtta 1260

ttgatcttta tttcaatcgt gacctaacag ctttggcgaa tgaacccgat gttgtcatca 1320

aaggagcatt caatggttgg aaatggaggc ttttcactga aagattgcat aagagtgacc 1380

ttggaggggt ttggtggtct tgcaaactgt acatacccaa ggaggcctac agattagact 1440

ttgtgttctt caacggtcgc acggtctatg agaacaatgg caacaatgat ttctgtatag 1500

gaatagaagg cactatgaat gaagatctgt ttgaggattt cttggttaaa gaaaagcaaa 1560

gggagcttga gaaacttgcc atggaagaag ctgaaaggag gacacagact gaagaacagc 1620

ggcgaagtaa ggaagcaagg gctgcagatg aagctgtcag ggcacaagcg aaggccgaga 1680

tagagatcaa gaacaaaaaa ttgcagagta tgttgagttt ggccagaaca tgtgttgata 1740

atttgtggta catagaggct agcacagata caagcggaga tactatcagg ttatactata 1800

acagaaactc gaggccactt gcgcatagta ctgagatttg gatgcatggt ggttacaaca 1860

attggtcaga tggactctct attgttgaaa gctttgtcaa gtgcaatgac agagacggcg 1920

attggtggta tgcagatggt acgacacctc aacctttgta cataaggcaa cattgttttg 1980

atttttttg ttgaggaaac atttgttttg attctagcat aatgctccta caaatatggc 2040

atgaatttcc ttgttttatt gatgtcatga gaaagtattt tattaactcg aaggccatgg 2100

aagctcaaca tttaccatag acagacgctt aaagatcatt tgtattccgt ggatcatata 2160

tgtaatgtaa tacctgtctt ttctctatat gtacagttat tccacctgaa aaagcacttg 2220

tgttggactg ggtttttgct gatgggccag ctgggaatgc aaggaactat gacaacaatg 2280

ctcgacaaga tttccatgct attcttccaa acaacaatgt aaccgaggaa ggcttctggg 2340

tgcaagagga gcaaaacatc tatacaaggc ttctgcaaga aaggagagaa aaggaagaaa 2400

ccatgaaaag aaaggtgagt tgcaacaaaa tctttgcata tagatc 2446

<210> 13
<211> 1032
<212> DNA
<213> Triticum sp.

<400> 13
gatctctata attttggcag ttaacccctg agtgatggca aatatattcc ctttcgtcta 60

ttttccaaat tcaaaatgca tggttccatg caagcttatc caaaatcact tgataatata 120

ccaatcacaa cataactttg tttaccataa gaacattcct acttaaaatt tgcaaggtaa 180

ctccctttcg aggctggttg gcttgatgag taactggcaa ttaacaaaga aaagatatat 240

ctgatgtttg gaacaaaaca tatgatcagg gttgtttggg ttgactcatg ttccttttta 300

cctacacagg ctgagagaag tgcaaatatc aaagctgaga tgaaggcaaa aactatgcga 360

aggtttctgc tttcccagaa acacattgtt tataccgaac cgcttgaaat acgtgccgga 420

accacagtgg atgtgctata caatccctct aacacagtgc taaatggaaa gccggaggtt 480

tggtttagat gctcttttaa cctttggatg catccaagtg gagcattgcc accccagaag 540

atggtgaaat caggggatgg gccgctctta aaagccacag gtttattgcg ttattacatc 600

actgttatta gtatatatat aaccattttt atgcaatcaa tagagtcaag tgcaactaat 660

gatgcacaga taggatcaca tcattaggag aatgatgtga tggacaagac ccaatcctaa 720

gcatagcaca agatcgtgta gttcgttcgc tagagctttt ctaatgtcaa gtatcatttc 780

cttagaccat gagattgtgc aactcccgga tatcgtagga gtgctttggg tgtatcaaat 840

gtcacaacgt aactgggtga ctataaaggt gcactacagg tatctccgaa agtttctgtt 900

gggttggcac gaatcgagac tgggatttgt cactccgtat gacggagagg tatctttggg 960

cccactcggt aatgcatcat cataatgagc tcaatgtgac taaggagtta gccacgggat 1020

cgagaattcc cg 1032

<210> 14
<211> 892
<212> DNA
<213> Triticum sp.

<400> 14
aatatttctt gttctattat tggtaataat tagctagttt aatgccataa gcccataaca 60

gatatgcaac tactccctcc aatccatatt acttgtcgca actttggtac aactttagta 120

caaagttata ctaaagctgt gacaagtaat atggaccgga gggagtacta tataagcttg 180

tagctgtttt gagaccgagt gtctgctcgg gtggctagct ggagcgggct gaagtgcttg 240

caggcacctc ttctctaaaa aaaagtgctt gcagcccccc cgccccctcc atagggtgag 300

tggtcacctt tcttcttaaa aattatggca ccaagggaaa ttctcggctg gtcgagcttg 360

tagctatttt ttcggagcgt gaatgggagc gtctttctgt ataaggccta taggcttact 420

ttgatatata ttgtgaagtc acttaagcct tgttaaaacg tagaaactta gttccgcaac 480

ttggccaaat ccctgttaaa ttggtttact gtgtactaga tgcatcgatg gcgcagagtc 540

ccggggggta ataaagcttc cattttctac aatgaagtta attatcctac ttgccttgta 600

attactgagt acaatacaga gcaccgaaaa gctgtatcct tcctacttcc ttatgtttat 660

ctgtgttcct tgtctagtta atgttccacc ggatgcctat atgatggact ttgttttctc 720

cgagtgggaa gaagatggga tctatgacaa caggaatggg atggactatc atattcctgt 780

ttctgattca attgaaacag agaattacat gcgtattatc cacattgccg ttgagatggc 840

ccccgttgca aaggtaatat aattctaagg ctagtttctt tgatgcgagg cg 892

<210> 15
<211> 871
<212> DNA
<213> Triticum sp.

<400> 15
aggttatcct ccagaatgaa ttttttccag tacgtattat ttagaatact agcggtatat 60

tgactttttc tttgtgagac tacactttct tgtttaccat tccagtgcac catgttcaaa 120

atcttgtatt cagcgcgtta ctttcagttt ctttactact agcttatttg gtgcattggt 180

gtttcctttc ctactctact atctgaatgc tacttgtgtt ttcgcaacag ttgcttcttt 240

atccccttcc atttctcagt taaaaaaact tgcatctgta ttcacgtgac agcatataat 300

acattgccat gattggtcaa gtgctccggt cgcctggcta tataaggaac actattccca 360

atccagaatg gcaagcactc gggttgtatt taccatccac aatcttgaat ttggagcaca 420

ttatattggt aaagcaatga catactgtga taaagccaca actgtgagtg ccttactgtc 480

ttgtaatttt taatctttct gtttggcgca cagaaaatct tccacatttt acagaatcat 540

gttcttgtgt tttgtacgta ttcaactatt tccacccaaa cttttcaggt ttctcctaca 600

tattcaaggg acgtggcagg ccatggtgcc attgctcctc atcgtgagaa attctacggc 660

attctcaatg gaattgatcc agatatctgg gatcctgatt gccaacatgc tgtttggtcg 720

tctcgaggtc tttacattgc tggtgctctt taccccgact ttctggcgtg aatgatggag 780

taatacgtga aaacattaat tcttttctca acaagggacg gacaaacgcg cgagattgcc 840

tcctacctgg cttcggaact gaaagaactg g 871

<210> 16
<211> 1592
<212> DNA
<213> Triticum sp.

<400> 16
cgggaattct cgatcccgtg gctaactcct tagtcacatt gagctcatta tgatgatgca 60

ttaccgagtg ggcccaaaga tacctctccg tcatacggag tgacaaatcc cagtctcgat 120

tcgtgccaac ccaacagaaa ctttcggaga tacctgtagt gcacctttat agtcacccag 180

ttacgttgtg acatttgata cacccaaagc actcctacga tatccgggag ttgcacaatc 240

tcatggtcta aggaaatgat acttgacatt agaaaagctc tagcgaacga actacacgat 300

cttgtgctat gcttaggatt gggtcttgtc catcacatca ttctcctaat gatgtgatcc 360

atacactgac aattttatcc cggtaccaga tttttcccca gagtgcaagt agatatatac 420

caaggccaca gatagtttta tgcttaacta tgtgtttcat actacttcag gtcccttata 480

cttgtgagaa tgttgtcgaa ggcaagagag ctgcaaaaag ggccttgcag cagaagtttg 540

gattacagca aactgatgtc cctattgtcg gaatcatcac ccgtctgaca gcccagaagg 600

gaatccacct catcaagcac gcaattcacc gaaccctcga aagcaacgga caggttcatc 660

atcccttgtg aacgaataaa catcaaacgt tttgtttata aaaagttgct tactatttgt 720

ttttgtttac ttcaaaacaa aagtctgaaa atgaagtgtt tggttcctag gtggttttgc 780

ttggttcagc tccagatcat cgaatacaag gcgatttttg cagattggcc gatgctcttc 840

acggtgttta ccacggtagg gtgaagcttg ttctaaccta cgatgagcct ctttctcacc 900

tggtgagctc caatatccta cacaccatct agccagccct tcattatggg agctggagac 960

tactttataa tttaggttga tgatcgatca tgctgcagat atacgctggc tccgacttca 1020

ttattgtccc ttcaatcttc gaaccctgtg gcttaacaca acttgttgcc atgcgttatg 1080

gatcgatccc tatagttcgg aaaaccggag gtgtgtgact atttctctcc attatgctgc 1140

actgatttgc atatgtcgag ctgttggaca tgaaatggaa actatccttt ggtatcgcag 1200

gactttacga cactgtcttc gacgtagaca atgataagga ccgggctcgg tctcttggtc 1260

ttgaaccaaa tgggttcagt ttcgacggag ccgacagcaa cggcgtggat tatgccctca 1320

acaggcaagt atcgttcctc aattagccct gaattcagca gtagtgctag gttatttacc 1380

ttgcatgttc catacctcat ttcagagcaa tcggcgcttg gttcgatgcc cgtgattggt 1440

tccactccct gtgtaagagg gtcatggaac aagactggtc atggaaccgg cccgcactgg 1500

actacattga attgtaccat gccgctcgaa aattctgaca cccaactgaa ccaatggcaa 1560

gaacaagcgc attgtgggat cgagaattcc cg 1592

<210> 17
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:PEPTIDE MOTIF

<400> 17
Asp Val Gln Leu Val Met Leu Gly Thr Gly
1 5 10

<210> 18
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:PEPTIDE MOTIF

<400> 18
Ala Ala Gly Lys Lys Asp Ala Gly Ile Asp
1 5 10

<210> 19
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:PEPTIDE MOTIF

<400> 19
Ala Thr Gly Lys Lys Asp Ala Gly Ile Asp
1 5 10

<210> 20
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:PEPTIDE MOTIF

<400> 20
Ala Leu Gly Lys Lys Asp Ala Gly Ile Asp
1 5 10

<210> 21
<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:PEPTIDE MOTIF

<400> 21
Ala Thr Gly Lys Lys Asp Ala Leu
1 5

<210> 22
<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:PEPTIDE MOTIF

<400> 22
Ala Leu Gly Lys Lys Asp Ala Leu
1 5

<210> 23
<211> 13
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:PEPTIDE MOTIF

<400> 23
Ala Ala Gly Lys Lys Asp Ala Arg Val Asp Asp Ala Ala
1 5 10

<210> 24
<211> 13
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:PEPTIDE MOTIF

<400> 24
Ala Leu Gly Lys Lys Asp Ala Gly Ile Val Asp Gly Ala
1   5   10

<210> 25
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:PRIMER

<400> 25
tgttgaggtt ccatggcacg ttc 23

<210> 26
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:PRIMER

<400> 26
agtcgttctg ccgtatgatg tcg 23

<210> 27
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:PRIMER

<400> 27
ccaagtacca gtggtgaacg c 21

<210> 28
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:PRIMER

<400> 28
cggtgggatc caacggccc 19

<210> 29
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:PRIMER

<400> 29

ggaggtcttg gtgatgttgt 20

<210> 30
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:PRIMER

<400> 30
cttgaccaat catggcaatg 20

<210> 31
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:PRIMER

<400> 31
cattgccatg attggtcaag 20

<210> 32
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:PRIMER

<400> 32
accacctgtc cgttccgttg c 21

<210> 33
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:PRIMER

<400> 33
gcacggtcta tgagaacaat ggc 23

<210> 34
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:PRIMER

<400> 34
tctgcatacc accaatcgcc g 21

<210> 35
<211> 25

<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:PEPTIDE MOTIF

<400> 35
Lys Val Gly Gly Leu Gly Asp Val Val Thr Ser Leu Ser Arg Ala Val
1 5 10 15

Gln Asp Leu Gly His Asn Val Glu Val
20 25

<210> 36
<211> 25
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:PEPTIDE MOTIF

<400> 36
Lys Val Gly Gly Leu Gly Asp Val Val Thr Ser Leu Ser Arg Ala Ile
1 5 10 15

Gln Asp Leu Gly His Thr Val Glu Val
20 25

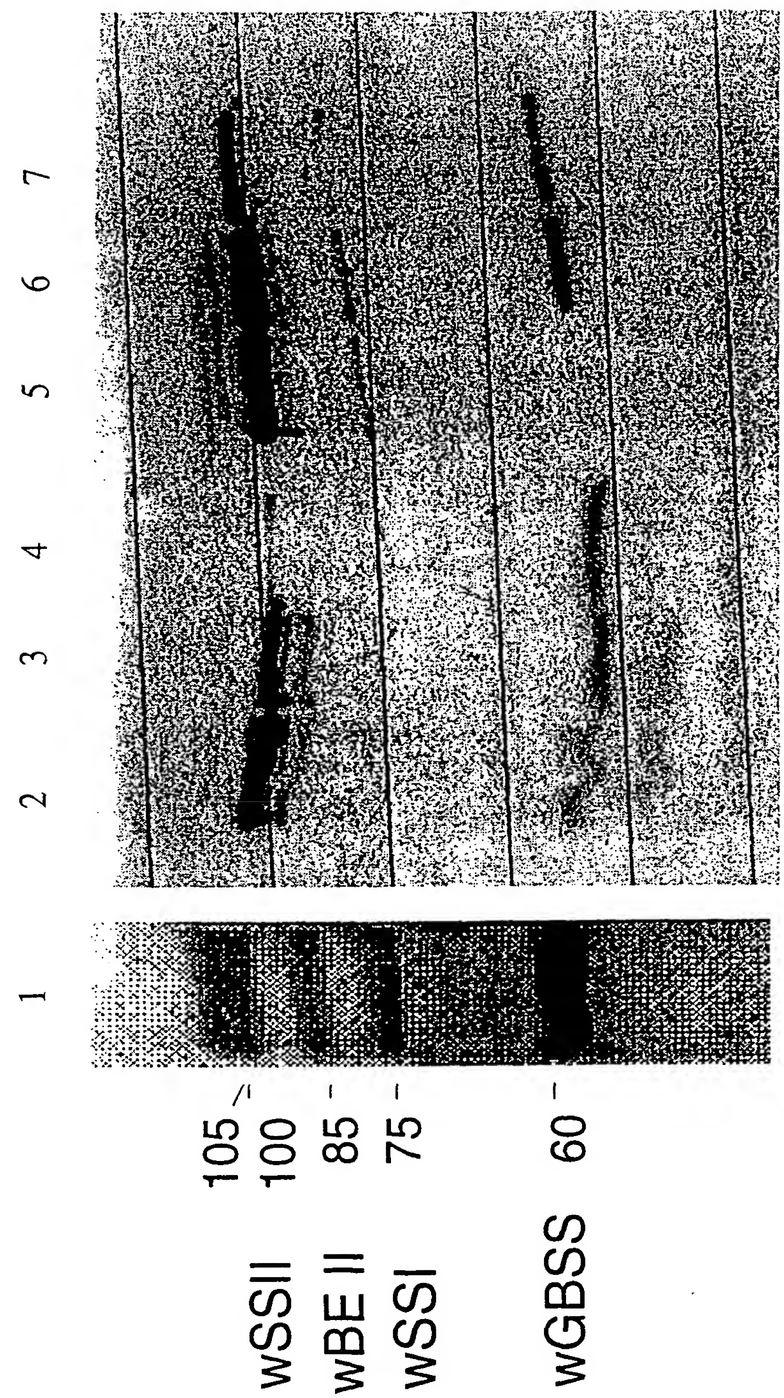


FIGURE 1

```
        1                                                   50
wSSIIB  ATTTCCTCGG CCTGACCCCG TGCGTTTACC CCACACAGAG CACACTCCAG
wSSIID  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
wSSIIA  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~

        51                                                 100
wSSIIB  TCCAGTCCAG CCCACTGCCG CGCTACTCCC CACTCCCACT GCCACCACCT
wSSIID  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
wSSIIA  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~GCT GCCACCACCT

        101                                                150
wSSIIB  CCGCcTGCGC CGCGCTCTGG GCGGACCAAC CCGCGCATCG TATCACGATC
wSSIID  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
wSSIIA  CCGCCTGCGC CGCGCTCTGG GCGGAGGACC AACCCGCGCA TCGTACCATC

        151                                                200
wSSIIB  ACCCACCCCG ATCCCGGCCG CCGCCATGTC GTCGGCGGTC GCGTCCGCCG
wSSIID  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
wSSIIA  GCCCGCCCCG ATCCCGGCCG CCGCCATGTC GTCGGCGGTC GCGTCCGCCG

        201                                                250
wSSIIB  CGTCCTTCCT CGCGCTCGCG TCCGCCTCCC CCGGGAGATC ACGGAGGAGG
wSSIID  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
wSSIIA  CGTCCTTCCT CGCGCTCGCC TCCGCCTCCC CCGGGAGATC ACGCAGGCGG

        251                                                300
wSSIIB  ACGAGGGTGA GCGCGTCGCC ACCCCACACC GGGGCTGGCA GGTTGCACTG
wSSIID  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
wSSIIA  GCGAGGGTGA GCGCGCCGCC ACCCCACGCC GGGGCCGGCA GGCTGCACTG

        301                                                350
wSSIIB  GCCGCCGTCG CCGCCGCAGC GCACGGCTCG CGACGGAGCG GTGGCCGCGC
wSSIID  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
wSSIIA  GCCGCCGTGG CCGCCGCAGC GCACGGCTCG CGACGGAGGT GTGGCCGCGC

        351                                                400
wSSIIB  GCGCCGCCGG GAAGAAGGAC GCGGGGAT.. .CGACGACGC CGCGCCCGCG
wSSIID  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
wSSIIA  GCGCCGCCGG GAAGAAGGAC GCGAGGGTCG ACGACGACGC CGCGTCCGCG

        401                                                450
wSSIIB  AGGCAGCCCC GCGCACTCCG CGGTGGCGCC GCCACCAAGG TTGCGGAGCG
wSSIID  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
wSSIIA  AGGCAGCCCC GCGCACGCCG CGGTGGCGCC GcCACCAAGG TCGCGGAGCG

        451                                                500
wSSIIB  GAGGGATCCC GTCAAGACGC TCGATCGCGA CGCCGCGGAA GGTGGCGCGC
wSSIID  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
wSSIIA  GAGGGATCCC GTCAAGACGC TCGATCGCGA CGCCGCGGAA GGTGGCGCGC

        501                                                550
wSSIIB  CGTCCCCGCC GGCACCGAGG CAGGAGGACG CCCGTCTGCC GAGCATGAAC
wSSIID  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
wSSIIA  CGGCACCGCC GGCACCGAGG CAGGACGCCG CCCGTCCaCC GAGTATGAAC
```

FIGURE 2-1

```
         551                                                    600
wSSIIB   GGCATGCCGG TGAACGGTGA AAACAAATCT ACCGGCGGCG GCGGCGCGAC
wSSIID   ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
wSSIIA   GGCACGCCGG TGAACGGTGA GAACAAATCT ACCGGCGGCG GCGGCGCGAC

         601                                                    650
wSSIIB   TAAAGACAGC GGGCTGCCCG CACCCGCACG CGCGCCCCAG CCGTCGAGCC
wSSIID   ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
wSSIIA   cAAAGACAGC GGGCTgcCCG CACCCGcACG CGCGCCCCAT cCGTCGAcCC

         651                                                    700
wSSIIB   AGAACAGAGT ACCGGTGAAT GGTGAAAACA AAGCTAACGT CGCCTCGCCG
wSSIID   ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
wSSIIA   AgAACAgAGT ACCAGTGAAC GGTGAAAACA AAGCTAACGT CGCCTCGCCG

         701                                                    750
wSSIIB   CCGACGAGCA TAGCCGAGGT CGCGGCTCCG GATCCCGCAG CTACCATTTC
wSSIID   ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
wSSIIA   CCGACGAGCA TAGCCGAGGT CGTGGCTCCG GATTCCGCAG CTACCATTTC

         751                                                    800
wSSIIB   CATCAGTGAC AAGGCGCCAG AGTCCGTTGT CCCAGCCGAG AAGGcgccgc
wSSIID   ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~CCAGCTGAG AAGACGCCGC
wSSIIA   CATCAGTGAC AAGGCGCCGG AGTCCGTTGT CCCAGCCGAG AAGCCGCCGC

         801                                                    850
wSSIIB   CGtCgtcCgg CtcAAATtTc gtgCcCtCgg cttctGctCc cggGtctGAC
wSSIID   CGTCGTCCGG CTCAAATTTC GAGTCCTCGG CCTCTGCTCC CGGGTCTGAC
wSSIIA   CGTCGTCCGG CTCAAATTTC GTGgTCTCGG CTTCTGCTCC CAGGCTGGAC

         851                                                    900
wSSIIB   actgtCaGCG acGtGGaact TgaActGAAg aAGGGtgCgg tCattgTcaA
wSSIID   ACTGTCAGCG ACGTGGAACA AGAACTGAAG AAGGGTGCGG TCGTTGTCGA
wSSIIA   ATTGACAGCG ATGTTGAACC TGAACTGAAG AAGGGTGCGG TCATCGTCGA

         901                                                    950
wSSIIB   aGAAgcTcCa aaCcCaAaGG CTCTTTCGCC GCCCGCAGCA CCCGCTGTAC
wSSIID   AGAAGCTCCA AAGCCAAAGG CTCTTTCGCC GCCtGCAGCc CCCGCTGTAC
wSSIIA   AGAAGCTCCA AACCCAAAGG CTCTTTCGCC GCCTGCAGCC CCCGCTGTAC

         951                                                   1000
wSSIIB   AACAAGACCT TTGGGACTTC AAGAAATACA TTGGTTTCGA GGAGCCCGTG
wSSIID   AAgAAGACCT TTGGGAtTTC AAGAAATACA TTGGTTTCGA GGAGCCCGTG
wSSIIA   AAGAAGACCT TTGGGACTTC AAGAAATACA TTGGCTTCGA GGAGCCCGTG

         1001                                                  1050
wSSIIB   GAGGCCAAGG ATGATGGCCG GGCTGTTGCA GATGATGCGG GCTCCTTCGA
wSSIID   GAGGCCAAGG ATGATGGCCG GGCTGTcGCA GATGATGCGG GCTCCTTtGA
wSSIIA   GAGGCCAAGG ATGATGGCTG GGCTGTTGCA GATGATGCGG GCTCCTTTGA

         1051                                                  1100
wSSIIB   ACACCACCAG AATCACGATT CCGGGCCTTT GGCAGGGGAG AACGTCATGA
wSSIID   ACACCACCAG AATCACGAcT CCGGaCCTTT GGCAGGGGAG AAtGTCATGA
wSSIIA   ACATCACCAG AACCATGATT CCGGACCTTT GGCAGGGGAG AACGTCATGA
```

FIGURE 2-2

```
         1101                                                1150
wSSIIB   ACGTGGTCGT CGTGGCTGCT GAATGTTCTC CCTGGTGCAA AACAGGTGGT
wSSIID   ACGTGGTCGT CGTGGCTGCT GAgTGTTCTC CCTGGTGCAA AACAGGTGGT
wSSIIA   ACGTGGTCGT CGTGGCTGCT GAATGTTCTC CCTGGTGCAA AACAGGTGGT

         1151                                                1200
wSSIIB   CTTGGAGATG TTGCCGGTGC TTTGCCCAAG GCTTTGGCGA AGAGAGGACA
wSSIID   CTgGGAGATG TTGCgGGTGC TcTGCCCAAG GCTTTGGCaA AGAGAGGACA
wSSIIA   CTTGGAGATG TTGCCGGTGC TTTGCCCAAG GCTTTGGCGA AGAGAGGACA

         1201                                                1250
wSSIIB   TCGTGTTATG GTTGTGGTAC CAAGGTATGG GGACTATGAG GAAGCCTACG
wSSIID   TCGTGTTATG GTTGTGGTAC CAAGGTATGG GGACTATGAa GAACCTACGg
wSSIIA   TCGTGTTATG GTTGTGGTAC CAAGGTATGG GGACTATGAG GAAGCCTACG

         1251                                                1300
wSSIIB   ATGTCGGAGT CCGAAAATAC TACAAGGCTG CTGGACAGGA TATGGAAGTG
wSSIID   ATGTCGGAGT CCGAAAATAC TACAAGGCTG CTGGACAGGA TATGGAAGTG
wSSIIA   ATGTCGGAGT CCGAAAATAC TACAAGGCTG CTGGACAGGA TATGGAAGTG

         1301                                                1350
wSSIIB   AATTATTTCC ATGCTTATAT CGATGGAGTT GATTTTGTGT TCATTGACGC
wSSIID   AATTATTTCC ATGCTTaTAT CGATGGAGTT GATTTTGTGT TCATTGACGC
wSSIIA   AATTATTTCC ATGCTTATAT CGATGGAGTT GATTTTGTGT TCATTGACGC

         1351                                                1400
wSSIIB   TCCTCTCTTC CGACACCGCC AGGAAGACAT TTATGGGGGC AGCAGACAGG
wSSIID   TCCTCTCTTC CGACACCGAG AGGAAGACAT TTATGGGGGC AGCAGACAGG
wSSIIA   TCCTCTCTTC CGACACCGCC AGGAAGACAT TTATGGGGGC AGCAGACAGG

         1401                                                1450
wSSIIB   AAATTATGAA GCGCATGATT TTGTTCTGCA AGGCCGCTGT CGAGGTTCCA
wSSIID   AAATTATGAA GCGCATGATT TTGTTCTGCA AGGCCGCTGT TGAGGTTCCA
wSSIIA   AAATTATGAA GCGCATGATT TTGTTCTGCA AGGCCGCTGT CGAGGTTCCT

         1451                                                1500
wSSIIB   TGGCACGTTC CATGCGGCGG TGTCCCTTAT GGGGATGGAA ATCTGGTGTT
wSSIID   TGGCACGTTC CATGCGGCGG TGTCCCTTAT GGGGATGGAA ATCTGGTGTT
wSSIIA   TGGCACGTTC CATGCGGCGG TGTCCCTTAT GGGGATGGAA ATCTGGTGTT

         1501                                                1550
wSSIIB   TATTGCAAAT GATTGGCACA CGGCACTCCT GCCTGTCTAT CTGAAAGCAT
wSSIID   TATTGCAAAT GATTGGCACA CGGCACTCCT GCCTGTCTAT CTGAAAGCAT
wSSIIA   TATTGCAAAT GATTGGCACA CGGCACTCCT GCCTGTCTAT CTGAAAGCAT

         1551                                                1600
wSSIIB   ATTACAGGGA CCATGGTTTG ATGCAGTACA CTCGGTCCAT TATGGTGATA
wSSIID   ATTACAGGGA CCATGGTTTG ATGCAGTACA CTCGGTCCAT TATGGTGATA
wSSIIA   ATTACAGGGA CCATGGTTTG ATGCAGTACA CTCGGTCCAT TATGGTGATA

         1601                                                1650
wSSIIB   CATAACATCG CTCACCAGGG CCGTGGCCCA GTAGATGAGT TCCCGTTCAC
wSSIID   CATAACATCG CTCACCAGGG CCGTGGCCCT GTAGATGAAT TCCCGTTCAC
wSSIIA   CATAACATCG CGCACCAGGG CCGTGGCCCA GTAGATGAAT TCCCGTTCAC
```

FIGURE 2-3

```
        1651                                                  1700
wSSIIB  CGAGTTGCCT GAGCACTACC TGGAACACTT CAGACTGTAC GACCCCGTGG
wSSIID  CGAGTTGCCT GAGCACTACC TGGAACACTT CAGACTGTAC GACCCCGTGG
wSSIIA  CGAGTTGCCT GAGCACTACC TGGAACACTT CAGACTGTAC GACCCCGTGG

        1701                                                  1750
wSSIIB  GTGGTGAACA CGCCAACTAC TTCGCCGCCG GCCTGAAGAT GGCGGACCAG
wSSIID  GTGGTGAACA CGCCAACTAC TTCGCCGCCG GCCTGAAGAT GGCGGACCAG
wSSIIA  GTGGTGAGCA CGCCAACTAC TTCGCCGCCG GCCTgAAGAT GgCGGACCAG

        1751                                                  1800
wSSIIB  GTTGTCGTCG TGAGCCCGGG GTACCTGTGG GAGCTGAAGA CGGTGGAGGG
wSSIID  GTTGTCGTGG TGAGCCCCGG GTACCTGTGG GAGCTGAAGA CGGTGGAGGG
wSSIIA  GTTGTCGTGG TGAGCCCCGG GTACCTGTGG gAGCTCAAGA CGGTGGAgGG

        1801                                                  1850
wSSIIB  CGGCTGGGGG CTTCACGACA TCATACGGCA GAACGACTGG AAGACCCGCG
wSSIID  CGGCTGGGGG CTTCACGACA TCATACGGCA GAACGACTGG AAGACCCGCG
wSSIIA  CGGCTGGGGG CTTCACGACA TCATACGGCA GAACGACTGG AAGACCCGCG

        1851                                                  1900
wSSIIB  GCATCGTGAA CGGCATCGAC AACATGGAGT GGAACCCCGA GGTGGACGTC
wSSIID  GCATCGTCAA CGGCATCGAC AACATGGAGT GGAACCCCGA GGTGGACGCC
wSSIIA  GCATCGTCAA CGGCATCGAC AACATGGAGT GGAACCCCGA GGTGGACGTC

        1901                                                  1950
wSSIIB  CACCTCAAGT CGGACGGCTA CACCAACTTC TCCCTGGGGA CGCTGGACTC
wSSIID  CACCTCAAGT CGGACGGCTA CACCAACTTC TCCCTGAGGA CGCTGGACTC
wSSIIA  CACCTCAAGT CGGACGGCTA CACCAACTTC TCCCTGGGGA CGCTGGACTC

        1951                                                  2000
wSSIIB  CGGCAAGCGG CAGTGCAAGG AGGCCCTGCA GCGGGAGCTG GGCCTGCAGG
wSSIID  CGGCAAGCGG CAGTGCAAGG AGGCCCTGCA GCGCGAGCTG GGCCTGCAGG
wSSIIA  CGGCAAGCGG CAGTGCAAGG AGGCCCTGCA GCGCGAGCTG GGCCTGCAGG

        2001                                                  2050
wSSIIB  TCCGCGGCGA CGTGCCGCTG CTCGGCTTCA TCGGGCGCCT GGACGGGCAG
wSSIID  TCCGCGCCGA CGTGCCGCTG CTCGGCTTCA TCGGCCGCCT GGACGGGCAG
wSSIIA  TCCGCGCCGA CGTGCCGCTG CTCGGCTTCA TCGGCCGCCT GGACGGGCAG

        2051                                                  2100
wSSIIB  AAGGGCGTGG AGATCATCGC GGACGCGATG CCCTGGATCG TGAGCCAGGA
wSSIID  AAGGGCGTGG AGATCATCGC GGACGCCATG CCCTGGATCG TGAGCCAGGA
wSSIIA  AAGGGCGTGG AGATCATCGC GGACGCCATG CCCTGGaTCG TGAGCCAGgA

        2101                                                  2150
wSSIIB  CGTGCAGCTG GTCATGCTGG GCACCGGGCG CCACGACCTG GAGGGCATGC
wSSIID  CGTGCAGCTG GTGATGCTGG GCACCGGGCG CCACGACCTG GAGAGCATGC
wSSIIA  CGTGCAGCTG GTCATGCTGG GCACCGGCCG CCACGACcTG gAGAGCATGc

        2151                                                  2200
wSSIIB  TGCGGCACTT CGAGCGGGAG CACCACGACA AGGTGCGCGG GTGGGTGGGG
wSSIID  TGCAGCACTT CGAGCGGGAG CACCACGACA AGGTGCGCGG GTGGGTGGGG
wSSIIA  TgCGGCACTT CGAGCGGGAG CACCACGACA AGGTGCGCGG gTGGGTGGGG
```

FIGURE 2-4

```
        2201                                               2250
wSSIIB  TTCTCCGTGC GGCTGGCGCA CCGGATCACG GCCGGCGCCG ACGCGCTCCT
wSSIID  TTCTCCGTGC GCCTGGCGCA CCGGATCACG GCGGGGGCGG ACGCGCTCCT
wSSIIA  TTCTCCGTgc GccTGGCGCA CCGGATCACG GCGGGCGCCG ACGCGCTCcT

        2251                                               2300
wSSIIB  CATGCCCTCC CGGTTCGAGC CGTGCGGACT GAACCAGCTC TACGCCATGG
wSSIID  CATGCCCTCC CGGTTCGTGC CGTGCGGGCT GAACCAGCTC TACGCCATGG
wSSIIA  CATGCCCTCC CGGTTCGAgC CGTGCGGGTT GAACCAGCTt TACGCCATGG

        2301                                               2350
wSSIIB  CCTACGGCAC CGTCCCCGTC GTGCATGCCG TCGGTGGCCT GAGGGACACC
wSSIID  CCTACGGCAC CGTCCCCGTC GTGCACGCCG TCGGCGGCCT CAGGGACACC
wSSIIA  CCTACGGCAC CGTCCCCGTC GTGCACGCCG TCGGCGGGGT GAGGGACACC

        2351                                               2400
wSSIIB  GTGCCGCCGT TCGACCCCTT CAACCACTCC GGGCTCGGGT GGACGTTCGA
wSSIID  GTGCCGCCGT TCGACCCCTT CAACCACTCC GGGCTCGGGT GGACGTTCGA
wSSIIA  GTGCCGCCGT TCGACCCCTT CAACCACTCC GGcCTCGGGT GGACGTTCGA

        2401                                               2450
wSSIIB  CCGCGCAGAG GCGCAGAAGC TGATCGAGGC GCTCGGGCAC TGCCTCCGCA
wSSIID  CCGCGCCGAG GCGCACAAGC TGATCGAGGC GCTCGGGCAC TGCCTCCGCA
wSSIIA  CCGCGCcGAG GCGCAcAAGC TGATCGAGGC GCTCGGGCAC TGCCTCCGCA

        2451                                               2500
wSSIIB  CCTACCGGGA CTACAAGGAG AGCTGGAGGG GGCTCCAGGA GCGCGGCATG
wSSIID  CCTACCGAGA CTTCAAGGAG AGCTGGAGGG CCCTCCAGGA GCGCGGCATG
wSSIIA  CCTACCGGGA CTACAAGGAG AGCTGGAGGG GcCTCCAGGA GCGCGGCATG

        2501                                               2550
wSSIIB  TCGCAGGACT TCAGCTGGGA GCATGCCGCC AAGCTCTACG AGGACGTCCT
wSSIID  TCGCAGGACT TCAGCTGGGA GCACGCCGCC AAGCTCTACG AGGACGTCCT
wSSIIA  TCGCAGGACT TCAGCTGGGA GCATGCCGCC AAGCTCTACG AGGACGTCCT

        2551                                               2600
wSSIIB  CGTCAAGGCC AAGTACCAGT GGTGAACGCT AGCTGCTAGC CGGTCCAGCC
wSSIID  CGTCAAGGCC AAGTACCAGT GGTGAACGCT AGCTGCTAGC CGCTCCAGCC
wSSIIA  CcTCAAGGCC AAGTACCAGT GGTGAACGCT AGCTGCTAGC CGcTCCAGCC

        2601                                               2650
wSSIIB  CCGCATGCG. ...TGCATGA CAGGATGGAA TTGCGCATTG CGCACGCAGG
wSSIID  CCGCATGCG. ...TGCATGA CAGGATGGAA CT..GCATTG CGCACGCAGG
wSSIIA  CCGCATGCGT GCATGcatgA gAGGgTGGAA cTGCGCATTG CGCcCGCAGG

        2651                                               2700
wSSIIB  AAGGTGCCAT .......... .GGAGCGCCG GCATCCGCGA AGTACAGTGA
wSSIID  AAAGTGCCAT .......... .GGAGCGCCG GCATCCGCGA AGTACAGTGA
wSSIIA  AAcGTGCCAT ccttctcgat gGGAGCGCCG GCATCCGCGA gGTgCAGTGA

        2701                                               2750
wSSIIB  CAT..GAGGT GTGTGTGGTT GAGACGCTGA TTC......C GATCTGGTCC
wSSIID  CAT..GAGGT GTGTGTGGTT GAGACGCTGA TTC......C AATCCGGCCC
wSSIIA  CATGAGagGT GTGTGTGGTT GAGACGCTGA TTCCGATCTc gatctGGTCC
```

FIGURE 2-5

```
         2751                                              2800
wSSIIB   GTAGCAGAGT AGAGCGGAGG TAGGGAAGCG CTCCTTGTTA CAGGTATATG
wSSIID   GTAGCAGAGT AGAGCGGAGG TATATGGGAA TCTTAACTTG GTATTGTAAT
wSSIIA   GTAGCAGAGT AGAGCGGAcG TAGGGAAGCG CTCCTTGTTg CAGGTATATG

         2801                                              2850
wSSIIB   GGAATGTTGT TAACTTGGTA TTGTAATTTG TTATGTTGTG TGCATTATTA
wSSIID   TTGTTATGTT GTGTGCATTA TTACAATGTT GTTACTTATT CTTGTTAAGT
wSSIIA   GGAATGTTGT cAACTTGGTA TTGTAgTTTG cTATGTTGTa TGCgTTATTA

         2851                                              2900
wSSIIB   CAGAGGGCAA CGATCTGCGC CGGCGCACCG GCCCAACTGT TGGGCCGGTC
wSSIID   CGGAGGCCAA GGGCGAAAGC TAGCTCACAT GTCTGATGGA TGCAAAAAAA
wSSIIA   caatgttgtt acttattctt gtTAAAAAAA AAAAAAAAAA AAAA~~~~~~

         2901                                              2950
wSSIIB   GCACAGCAGC CGTTGGATCC GACCGCCTGG GCCGTTGGAT CCCACCGAAA
wSSIID   AAAAAAAAAA AAA~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
wSSIIA   ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~

         2951       2965
wSSIIB   AAAAAAAAAA AAAAA
wSSIID   ~~~~~~~~~~ ~~~~~
wSSIIA   ~~~~~~~~~~ ~~~~~
```

FIGURE 2-6

```
WSSIIA      1 MSSAVASAAS ---FLALASA SP-GRSRRRA RVSAPPPHAG AGRL----HW PPWPP-QRTA  51
WSSIIB      1 ********** ---******* **-******T ****S***T* *******---** **S**-****  51
WSSIID        ---------- ---------- ---------- ---------- ---------- ----------
ZSSIIA      1 ****AV*SS* STF******* **G**--*** **GSS*F*T* *-S*SFAFWA **S**RAPRD  57
ZSSIIB      1 *PG*-I*SS* SAFL*PV**S **---R***G S*G*ALRSY* YSGAELRL** ARRG*P*DG*  56
PEASSII     1 *MLSLG*D*T VLP*H*KNLK FTPKL*TLNG --DLAFSKGL GVGRLNCGSV ---------R  49
POTSSII    10 PVNFIFCDFY VMENSI*LHS GNQFHPNLPL ---LALRPKK LSLIHGSSRE ---------Q  57

                  ⇓ Transit peptide cleavage site
WSSIIA.    52 RDGGVAARAA GKKDARVDDD AASARQPRAR RGGAATKVAE RRDPVKTLDR DAAEGGAPAP 111
WSSIIB     52 ***A****** *****GI-** **P******L ********** ********** ********S* 110
WSSIID        ---------- ---------- ---------- ---------- ---------- ----------
ZSSIIA     58 AALVR*EAE* *G***PPERS GDA**L**** *----NA*SK ****------ ----------  97
ZSSIIB     57 -ASVR**A*P AGG------- ---------- ---------- ---------- ----------  68
PEASSII    50 LNHKQHV**V **SFGADENG DG*EDDVVNA TIEKSK**LA LQRELIQQIA ERKKLVSSID 109
POTSSII    58 MWRNQRVK*T *ENSGEAA-S *DESNDALQV TIEKSK**LA MQQDLLQQIA ERRKVVSSIK 116

WSSIIA    112 PAPRQDAARP PSMNGTPVNG ENKSTGGGGA TKDSGLPAPA RAPHPSTQNR VPVNGENKAN 171
WSSIIB    111 *****ED**L *****M**** ********** ********** ***Q**S*** ********** 170
WSSIID        ---------- ---------- ---------- ---------- ---------- ----------
ZSSIIA     98 ---------- ---------- -----LQPVG RYG*ATGNT* *TGAA*C**A ALADV*I*SI 132
ZSSIIB     69 ---------- ---------- ---------- -ESEEAAKSS SSSQAGAVQG STAKAVDS*S  97
PEASSII   110 SDSIPGLEGN GVSYESSEKS LSR------- ---------- ------DS*P QKGSSSSGSA 146
POTSSII   117 S----SL*NA KGTYDGGSGS LSDVDIPDVD KDYNVTVPST A*TGITDVDK NTPPAISHDF 172

WSSIIA    172 VASPPTSIAE VVAPDSAATI SISDKAPESV VPAEKPPPSS GSNFVVSASA PRLDIDSDVE 231
WSSIIB    171 ********** *A***P**** ********** *****A**** *****P**** *GS*TV**** 230
WSSIID    203 ---------- ---------- ---------- -****T**** ****ES**** *GS*TV**** 231
ZSSIIA    134 **A*****VK FP**GYRMIL PSG*I***T* L**P**--LH E*PA*DGD*N --GIAPPT** 188
ZSSIIB     99 PPN*L**APK QSQSAAMQNG TSGGSSASTA A*VSG*KADH P*AP*TKREI DASAVKPEPA 158
PEASSII   147 *ETKR--WHC FQQ-----LC RSKETETWA* SSVGINQGFD EIEKKND*VK ASSKLHFNEQ 199
POTSSII   173 *E*KREIKRD LADERAPPLS RS*IT*SSQI SSTVSSK--R TL*VPPETPK SSQETLL**N 230

                                                         wSSIIp1 Region
WSSIIA    232 PELKKGAVIV EEAPNPKALS PPAAPAVQED LWDFKKYIGF EEPVEAKDDG WAVADDAGSF 291
WSSIIB    231 L********* K********* ********Q* ********** ********** R********* 290
WSSIID    232 Q*******V* ****K***** ********** ********** ********** R********* 291
ZSSIIA    189 *--------- ---------- -----L***A T********* D**D*****S RVG******* 224
ZSSIIB    159 GDDARPVESI ---------- ---------- --------*I A***D**A*- A*P*T**AAS 188
PEASSII   200 IKN*LYERPD TKDIS--SSI R--------- ------TSSL KFENFEGANE PSSKEV*NEA 242
POTSSII   231 SRKSLVD*PG KKIQSYMPSL R--------- ----*ESSAS HVEQRNENLE GSS*EANEET 277

                                            Region 1              Region 2
WSSIIA.   292 EHHQNHD--S GPLAGENVMN VVVVAAECSP WCKTGGLGDV AGALPKALAK RGHRVMVVVP 349
WSSIIB    291 *******--* ********** ********** ********** ********** ********** 348
WSSIID    292 *******--* ********** ********** ********** ********** ********** 349
ZSSIIA    225 **YGDN*--* ********** *I******** ********** V********R ********** 282
ZSSIIB    189 APYDRE*NEP *****P**** *****S**A* F********* V********R ********I* 248
PEASSII   243 *NFESGGEKP P****T**** IIL*S***A* *S******** **S******R ******I*A* 302
POTSSII   278 *DPV*I*EKP P****T**** IIL**S**A* *S******** *********R ********A* 337
```

FIGURE 3-1

```
                                            Sgp-1 Peptide 3
WSSIIA    350 RYGDYEEAYD VGVRKYYKAA GQDMEVNYFH AYIDGVDFVF IDAPLFRHRQ EDIYGGSRQE 409
WSSIIB    349 ********** ********** ********** ********** ********** ********** 408
WSSIID    350 *******PT* ********** ********** ********** *********E ********** 409
ZSSIIA    283 *****V**F* **I******* ***L****** *F******** ********** D********* 342
ZSSIIB    249 ***E*A**R* L***RR**V* ***S**T*** S********* VE**P****H NN****E*LD 308
PEASSII   303 H**N*A**H* I****R**V* ******T*** T******I** **S*I**NLE SN****N*LD 362
POTSSII   338 **DN*P*PQ* S****I**VD ***VD*T**Q *LLMDC**** *HSHM***IG NN****N*VD 397

                                                   Region 3
WSSIIA    410 IMKRMILFCK AAVEVPWHVP CGGVPYGDGN LVFIANDWHT ALLPVYLKAY YRDHGLMQYT 469
WSSIIB    409 ********** ********** ********** ********** ********** ********** 468
WSSIID    410 ********** ********** ********** ********** ********** ********** 469
ZSSIIA    343 ********** V********* ****C***** ********** ********** ********** 402
ZSSIIB    309 *L******** *******YA* ***TV***** ********** ********** ***N*****A 368
PEASSII   363 *LR**V**** ********** ***IC***** ********** ********** *******N** 422
POTSSII   398 *L***V**** **I******* ****C***** ********** ****A***** ***N*I*N** 457

WSSIIA    470 RSIMVIHNIA HQGRGPVDEF PFTELPEHYL EHFRLYDPVG GEHANYFAAG LKMADQVVVV 529
WSSIIB    469 ********** ********** ********** ********** ********** ********** 528
WSSIID    470 ********** ********** ********** ********** ********** ********** 529
ZSSIIA    404 **VL****** ********** *YMD****** Q**E****** *****I**** *****R**T* 462
ZSSIIB    369 **VL****** ********D* VNFD*****I D**K***NI* *D*S*V**** **T**R**T* 428
PEASSII   423 **VL****** *******ED* NTVD*SGN** DL*KM***** ***F*I**** **T**RI*T* 482
POTSSII   458 **VL****** ******LED* SYVD**P**M DP*K****** ***F*I**** **T**R**T* 517

                                           Region 4
WSSIIA    530 SPGYLWELKT VEGGWGLHDI IRQNDWKTRG IVNGIDNMEW NPEVDVHLK- SDGYTNFSLG 588
WSSIIB    529 ********** ********** ********** ********** *********- ********** 587
WSSIID    530 ********** ********** ********** ********** *****A***- *********R 588
ZSSIIA    463 *R******** ********** **S****IN* ******HQ** **K*****R- ******Y**E 521
ZSSIIB    429 *N**M***** S********* *N*****LQ* ******MS** **A*****H- **D***YTFE 487
PEASSII   483 *H**A***** S*******N* *NES***F** ****V*TKD* **QF*AY*T- ******YN*K 541
POTSSII   518 *H**S***** SQ******Q* *NE****LQ* ******TK** ***L****PR ****M*Y**D 577

                                   Region 5                      Region 5a
WSSIIA    589 TLDSGKRQCK EALQRELGLQ VRADVPLLGF IGRLDGQKGV EIIADAMPWI VSQDVQLVML 648
WSSIIB    588 ********** ********** **G******* ********** ********** ********** 647
WSSIID    588 ********** ********** ********** ********** ********** ********** 648
ZSSIIA    522 ***A****** A********E **D******* ********** D**G****** AG******** 581
ZSSIIB    488 ***T****** A****Q**** **D****I** *****H**** D*****IH** AG******** 547
PEASSII   542 **QT****** A********P **E***IIS* *****H**** DL**E*I**M M*H******* 601
POTSSII   578 **QT**P*** A***K****P **D****I** *****P**** DL**E*V**M MG******** 637

                                                      Region 6
WSSIIA    649 GTGRHDLESM LRHFEREHHD KVRGWVGFSV RLAHRITAGA DALLMPSRFE PCGLNQLYAM 708
WSSIIB    648 ********G* ********** ********** ********** ********** ********** 707
WSSIID    649 ********** *Q******** ********** ********** *********V ********** 708
ZSSIIA    582 ****A***R* *Q*L****PN ********** PM******** *V*V****** ********** 641
ZSSIIB    548 ****A***D* **R**S**S* ***A****** P********* *I******** ********** 607
PEASSII   602 ****A***Q* *KE**AQ*C* *I*S****** KM*******S *I******** ********** 661
POTSSII   638 ****R***Q* **Q**CQ*N* *I******** KTS******* *I******** **A******* 697
```

FIGURE 3-2

```
                 Region 7
WSSIIA   709 AYGTVPVVHA VGGVRDTVPP FDPFNHSGLG WTFDRAEAHK LTEALGHCLR TYRDYKESWR 768
WSSIIB   708 ********** ***L****** ********** ********Q* ********** ********** 767
WSSIID   709 ********** ***L****** ********** ********** ********** ****F***** 768
ZSSIIA   642 ********** ***L****A* ****GDA*** ********N* *****R***D ***K*G***K 701
ZSSIIB   608 ********** ***L****A* *****DT*** ********NR M*D**S***T ***N****** 667
PEASSII  662 S********G ***L****Q* *N**DE**V* ********N* *MA**WN**L **K***K**E 721
POTSSII  698 K***I***** ***L****Q* ***LMSQDW* GPS*****SQ **PRIRN**L ***E**K**E 757

WSSIIA   769 GLQERGMSQD FSWEHAAKLY EDVLLKAKYQ W 799
WSSIIB   768 ********** ********** ****V***** * 798
WSSIID   769 ********** ********** ****V***** * 799
ZSSIIA   702 S**A****** L**D***E** ****V***** * 732
ZSSIIB   668 ACRA***AE* L**D***V** ****V***** * 698
PEASSII  722 *I******** L**DN**QQ* *E**VA**** * 752
POTSSII  759 *I*T*C*T** L**DN**QN* *E**IA**** * 788
```

FIGURE 3-3

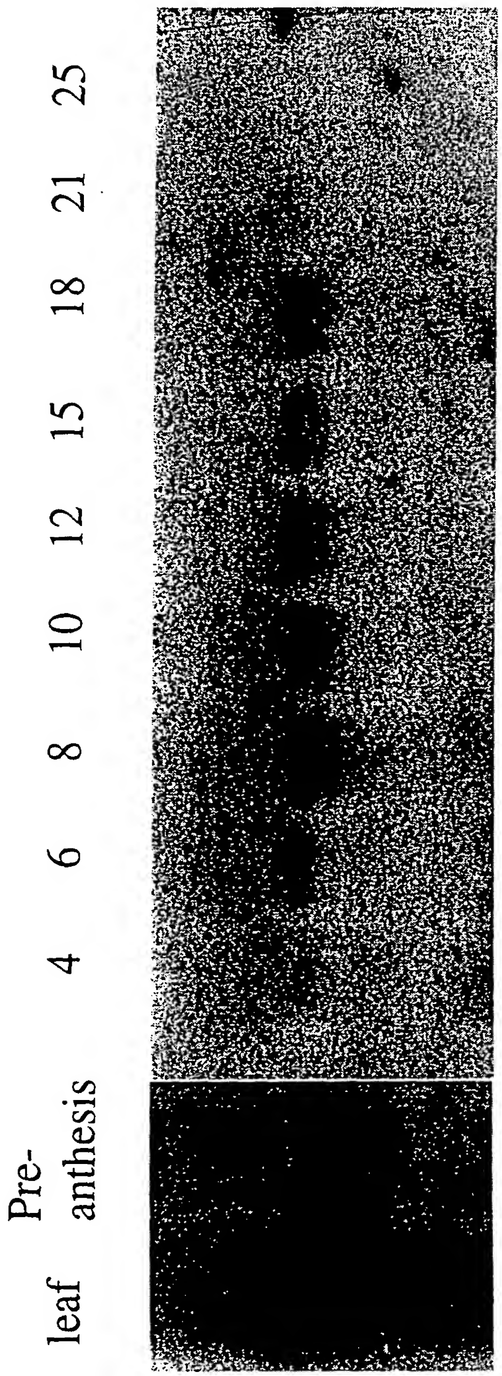


FIGURE 4

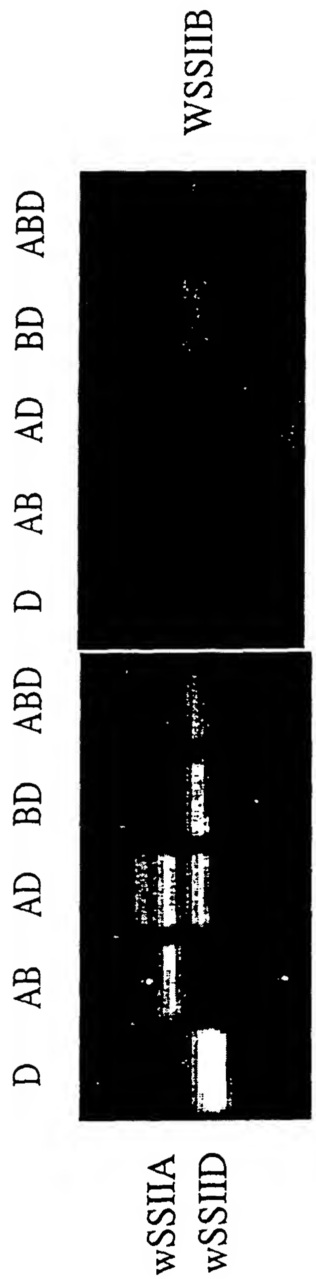


FIGURE 5

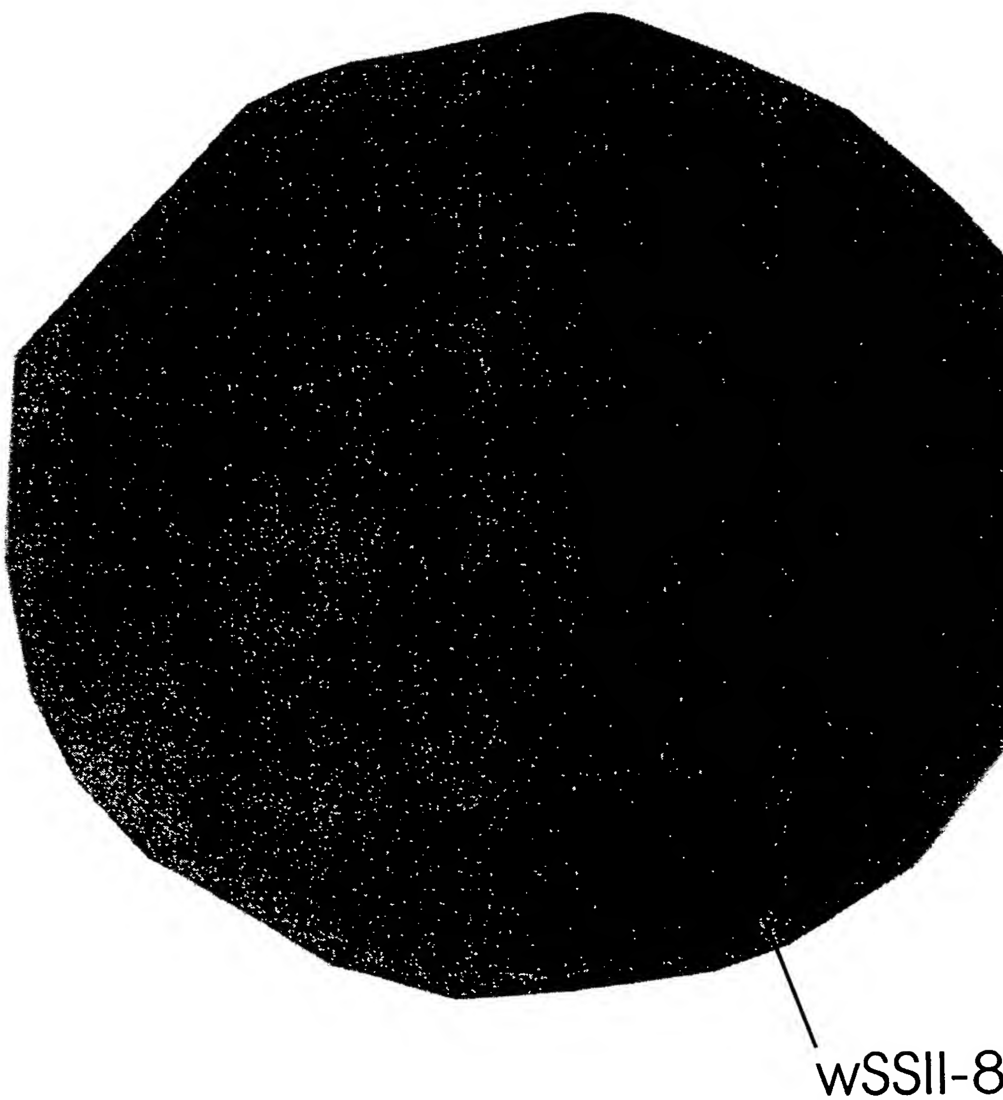


FIGURE 6

1 2 3 4 5 6 7 8 9 10 M

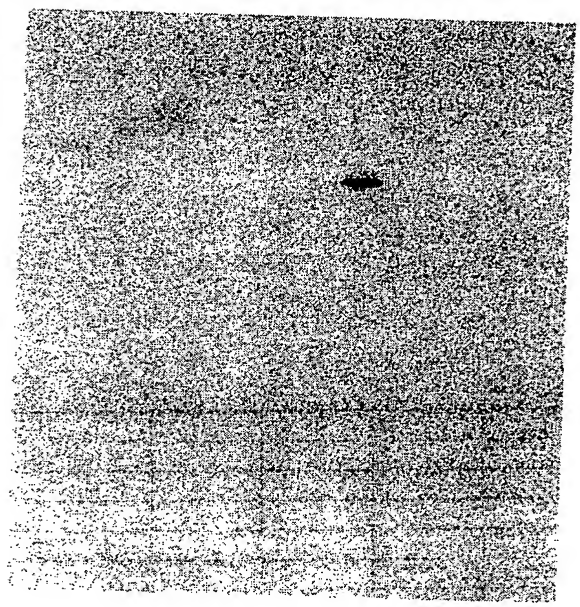


FIGURE 7

```
1                                                         50
 maizeSSIII   MEMVLRSQSP LCLRSGPVLI FRPTVAGGGG GTQSLLRTTR FARRRVIRCV
potatoSSIII   ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
 wheatSSIII   MEMSLWPRSP LCPRSRQPLV V.VRPAGRGG LTQPFLMNGR FTRSRTLRCM

              51                                                100
 maizeSSIII   VASPGCPNRK S.RTASPNVK VAAYSNYAPR LLVESSSKKS EHHDSSRHRE
potatoSSIII   ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
 wheatSSIII   VASSDPPNRK SRRMVPPQVK VISSRGYTTR LIVEPSNENT EHNNRD...E

              101                                               150
 maizeSSIII   ETIDTYNGLS GSDAAELTSN RDVEIEVDLQ HISEEELPGK VSINASLGEM
potatoSSIII   ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
 wheatSSIII   ETLDTYNALL STETAEWTDN REAE...... .TAKADSSQN ALSSSIIGEV

              151                                               200
 maizeSSIII   ETVDEAEVEE DKFEVDTSGI VLRNVAVREV DPKDEHNAKD VFVVDSSGTA
potatoSSIII   ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
 wheatSSIII   DVADEDILAA DLTVYSLSSV MKKEVDAADK ARVKE....D AFELDXASTT

              201                                               250
 maizeSSIII   PDNAAVEEVV DEAEVEEDMV DVDILGLDLN NATIEEIDLM EEALLENFDV
potatoSSIII   ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
 wheatSSIII   LRSVIVDVMD HXWDCQETLR SVIVDVMDHN GTVQETLRSV IVDVMDDAAD

              251                                               300
 maizeSSIII   DSPGNASSGR TYGGVDELGE LPSTSVDCIA INGKRRSLKP KPLPIVRFQE
potatoSSIII   ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
 wheatSSIII   KARVEEDVFE LDLSGNISSS ATTVELDAVD EVGPVQDKFE ATSSGNVSNS

              301                                               350
 maizeSSIII   QEQIVLSIVD EEGLIASSCE .EGQPVVDYD KQEENSTAFD EQKQLTDDFP
potatoSSIII   ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
 wheatSSIII   ATVREVDASD EAGNDQGIFR ADLSGNVFSS STTVEVGAVD EAGSIKDRFE

              351                                               400
 maizeSSIII   EEGISIVHFP EPNNDIVGSS KFLEQKQELD GSYKQDRSTT GLHEQDQSVV
potatoSSIII   ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
 wheatSSIII   TDSSGNVSTS APMWDAIDET VADQDTFEAD LSGNASSCAT YREVDDVVDE

              401                                               450
 maizeSSIII   SSHGQDKSIV GVPQQIQYND QSIAGSHRQD QSIAGAPEQI QSVAGYIKPN
potatoSSIII   ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~MDVPFPLH
 wheatSSIII   TRSEEETFAM DL.....FAS ESGHEKHMAV DYVGEATDEE ETYQQQYPVP

              451                                               500
 maizeSSIII   QSIVGSCKQH ELIIPEPKKI ESIISYNEID QSIVGSH.KQ DKSVVSVPEQ
potatoSSIII   RSLSCTSVSN AITHLKIKPI LGFVSHGTTS LSVQSSSWRK DGMVTGVSFS
 wheatSSIII   SSFSMWDKAI AKTGVSLNPE LRLVRVEE.. QGKVNFSDKK DLSIDDLPGQ

              501                                               550
 maizeSSIII   IQSIVSHSKP NQSTVDSYRQ AESIIGVPEK VQSITSYDKL DQSIVGSLKQ
potatoSSIII   ICANFSGRRR RKVSTPRSQG SSPKGFVPRK PSGMSTQRKV QKS.NGDKES
```

FIGURE 8-1

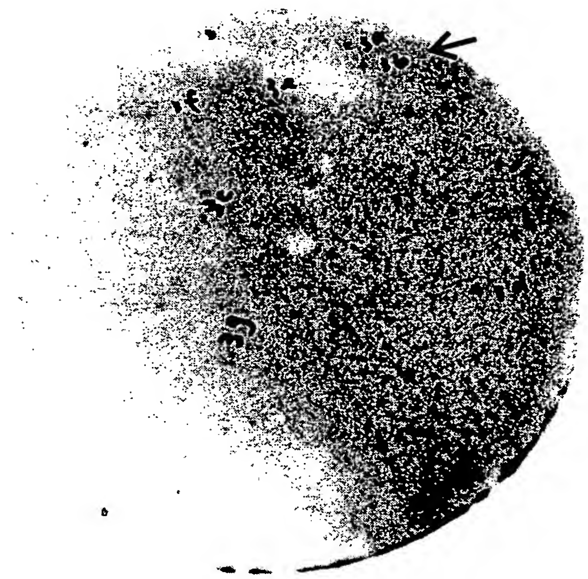
```
wheatSSIII   .......... NQSIIGSYKQ DKSIADVAGP TQSIFGSSKQ HRSIVAFPKQ

             551                                                600
maizeSSIII   DEPIISVPEK IQSIVHYTKP NQSIVGLPKQ QQSIVHIVEP KQSIDGFPKQ
potatoSSIII  KSTSTSKESE ISNQKTVEAR VETSDDDTKG VVRDHKFLED EDEINGSTKS
wheatSSIII   NQSIVSVTEQ KQSIVGFRSQ DLSAVSLPKQ NVPIVGYVER GSNXKQVPVV

             601                                                650
maizeSSIII   .DLSIVGISN EFQTKQLATV GTHDGLLMKG VEAKETSQKT EGDTLQATFN
potatoSSIII  ISMSPVRVSS QFVESEETGG DDKDAVKLNK SKRSEES... .GFIIDSVIR
wheatSSIII   DRQDALYVNG .......... .......LEA KEGDHTSEKT DEDALHVKFN

             651                                                700
maizeSSIII   VDNLSQKQEG LTKEADEITI IEKINDEDLV MIEEQKSIAM NEEQTIVTEE
potatoSSIII  EQSGSQGETN ASSKGSH.AV GTKLYEILQV DVEPQQ...L KENNAGNVEY
wheatSSIII   VDNVLRKHQA DRTQAVEKKT WKKVDEEHLY MTEHQKRAA. .EGQMVVNED

             701                                                750
maizeSSIII   DIPMAKVEIG IDKAKFL.HL LSEEESSWDE NEVGIIEADE QYEVDETSMS
potatoSSIII  KGPVASKLLE ITKASDVEHT ESNEIDDLDT NSFFKSDLIE EDEPLAAGTV
wheatSSIII   ELSIT..EIG MGRGDKIQHV LSEEELSWSE DEVQLIEDDG QYEVDETSVS

             751                                                800
maizeSSIII   ..TEQDIQES PNDDLDPQAL WSMLQELAEK NYSLGNKLFT YPDVLKADST
potatoSSIII  ..ETGDSSLN LRLEMEANLR RQAIERLAEE NLLQGIRLFC FPEVVKPDED
wheatSSIII   VNVEQDIQGS PQDVVDPQAL KVMLQELAEK NYSMRNKLFV FPEVVKADSV

             801                                                850
maizeSSIII   IDLYFNRDLS AVANEPDVLI KGAFNGWKWR FFTEKLHKSE LAGDWWCCKL
potatoSSIII  VEIFLNRGLS TLKNESDVLI MGAFNEWRYR SFTTRLTETH LNGDWWSCKI
wheatSSIII   IDLYLNRDLT ALANEPDVVI KGAFNGWKWR LFTERLHKSD LGGVWWSCKL

             851                                                900
maizeSSIII   YIPKQAYRMD FVFFNGHTVY ENNNNNDFVI QIESTMDENL FEDFLAEEKQ
potatoSSIII  HVPKEAYRAD FVFFNGQDVY DNNDGNDFSI TVKGGMQIID FENFLLEEKW
wheatSSIII   YIPKEAYRLD FVFFNGRTVY ENNGNNDFCI GIEGTMNEDL FEDFLVKEKQ

             901                                                950
maizeSSIII   RELENLANEE AERRRQTDEQ RRMEEERAAD KADRVQAKVE VETKKNKLCN
potatoSSIII  REQEKLAKEQ AERERLAEEQ RRIEAEKAEI EADRAQAKEE AAKKKKVLRE
wheatSSIII   RELEKLAMEE AERRTQTEEQ RRRKEARAAD EAVRAQAKAE IEIKKKKLQS

             951                                               1000
maizeSSIII   VLGLARAPVD NLWYIEPITT GQEATVRLYY NINSRPLVHS TEIWMHGGYN
potatoSSIII  LMVKATKTRD ITWYIEPSEF KCEDKVRLYY NKSSGPLSHA KDLWIHGGYN
wheatSSIII   MLSLARTCVD NLWYIEASTD TRGDTIRLYY NRNSRPLAHS TEIWMHGGYN

             1001                                              1050
maizeSSIII   NWIDGLSFAE RLVHHHDKDC DWWFADVVVP ERTYVLDWVF ADGPPGSARN
potatoSSIII  NWKDGLSIVK KLVKSERIDG DWWYTEVVIP DQALFLDWVF ADGPPKHAIA
wheatSSIII   NWXDGLSIVE SFVKCNDKDG DWWYADVIPP EKALVLDWVF ADGPAGNARN

             1051                                              1100
maizeSSIII   YDNNGGHDFH ATLP.NNMTE EEYWMEEEQR IYTRLQQERR EREEAIKRKA
potatoSSIII  YDNNHRQDFH AIVP.NHIPE ELYWVEEEHQ IFKTLQEERR LREAAMRAKV
wheatSSIII   YDNNARQDFH AILPNNNVTE EGFWAQEEQN IYTRLLQERR EKEETMKRKA

             1101                                              1150
```

FIGURE 8-2

```
 maizeSSIII   ERNAKMKAEM KEKTMRMFLV SQKHIVYTE. PLEIHAGTTI DVLYNPSNTV
potatoSSIII   EKTALLKTET KERTMKSFLL SQKHVVYTE. PLDIQAGSSV TVYYNPANTV
 wheatSSIII   ERSANIKAEM KAKTMRRFLL SQKHIVYTRT XLKYVPGTTV DVLYNPSNTV

              1151                                              1200
 maizeSSIII   LTGKPEVWFR CSFNRWMYPG GVLPPQKMVQ AENGSHLKAT VYVPRDAYMM
potatoSSIII   LNGKPEIWFR CSFNRWTHRL GPLPPQKMSP AENGTHVRAT VKVPLDAYMM
 wheatSSIII   LNGKSEGWFR CSFNLWMHSS GALPPQKMVK SGDGPLLKAT VDVPPDAYMM

              1201                                              1250
 maizeSSIII   DFVFSESEEG GIYDNRNGLD YHIPVFGSIA KEPPMHIVHI AVEMAPIAKV
potatoSSIII   DFVFSEREDG GIFDNKSGMD YHIPVFGGVA KEPPMHIVHI AVEMAPIAKV
 wheatSSIII   DFVFSEWEED GIYDNRNGMD YHIPVSDSIE TENYMRIIHI AVEMAPVAKV

              1251                                              1300
 maizeSSIII   GGLGDVVTSL SRAVQDLGHN VEVILPKYGC LNLSNVKNLQ IHQSFSWGGS
potatoSSIII   GGLGDVVTSL SRAVQDLNHN VDIILPKYDC LKMNNVKDFR FHKNYFWGGT
 wheatSSIII   GGLGDVVTSL SRAIQDLGHT VEVILPKYDC LNQSSVKDLH LYQSFSWGGT

              1301                                              1350
 maizeSSIII   EINVWRGLVE GLCVYFLEPQ NGMFGVGYVY G.RDDDRRFG FFCRSALEFL
potatoSSIII   EIKVWFGKVE GLSVYFLEPQ NGLFSKGCVY GCSNDGERFG FFCHAALEFL
 wheatSSIII   EIKVWVGRVE DLTVYFLEPQ NGMFGVGCVY G.RNDDRRFG FFCHSALEFI

              1351                                              1400
 maizeSSIII   LQSGSSPNII HCHDWSSAPV AWLHKENYAK SSLANARVVF TIHNLEFGAH
potatoSSIII   LQGGFSPDII HCHDWSSAPV AWLFKEQYTH YGLSKSRIVF TIHNLEFGAD
 wheatSSIII   LQNEFSPHII HCHDWSSAPV AWLYKEHYSQ SRMASTRVVF TIHNLEFGAH

              1401                                              1450
 maizeSSIII   HIGKAMRYCD KATTVSNTYS KEVSGHGAIV PHLGKFYGIL NGIDPDIWDP
potatoSSIII   LIGRAMTNAD KATTVSPTYS QEVSGNPVIA PHLHKFHGIV NGIDPDIWDP
 wheatSSIII   YIGKAMTYCD KATTVSPTYS RDVAGHGAIA PHREKFYGIL NGIDPDIWDP

              1451                                              1500
 maizeSSIII   YNDNFIPVHY TCENVVEGKR AAKRALQQKF GLQQIDVPVV GIVTRLTAQK
potatoSSIII   LNDKFIPIPY TSENVVEGKT AAKEALQRKL GLKQADLPLV GIITRLTHQK
 wheatSSIII   YTDNFIPVPY TCENVVEGKX AAKRALQQKF GLQQTDVPIV GIITRLTAQK

              1501                                              1550
 maizeSSIII   GIHLIKHAIH RTLERNGQVV LLGSAPDSRI QADFVNLANT LHGVNHGQVR
potatoSSIII   GIHLIKHAIW RTLERNGQVV LLGSAPDPRV QNNFVNLANQ LHSKYNDRAR
 wheatSSIII   GIHLIKHAIH RTLESNGQVV LLGSAPDHRI QGDFCRLADA LHGVYHGRVK

              1551                                              1600
 maizeSSIII   LSLTYDEPLS HLIYAGSDFI LVPSIFEPCG LTQLVAMRYG TIPIVRKTGG
potatoSSIII   LCLTYDEPLS HLIYAGADFI LVPSIFEPCG LTQLTAMRYG SIPVVRKTGG
 wheatSSIII   LVLTYDEPLS HLIYAGSDFI IVPSIFEPCG LTQLVAMRYG SIPIVRKTGG

              1601                                              1650
 maizeSSIII   LFDTVFDVDN DKERARDRGL EPNGFSFDGA DSNGVDYALN RAISAWFDAR
potatoSSIII   LYDTVFDVDH DKERAQQCGL EPNGFSFDGA DAGGVDYALN RALSAWYDGR
 wheatSSIII   LXDTVFDVDN DKDRARSLGL EPNGFSFDGA DSNGVDYALN RAIGAWFDAR

              1651                                   1686
 maizeSSIII   SWFHSLCKRV MEQDWSWNRP ALDYIELYRS ASKL*~
potatoSSIII   DWFNSLCKQV MEQDWSWNRP ALDYLELYHA ARKLE*
 wheatSSIII   DWFHSLCKRV MEQDWSWNRP ALDYIELYHA ARKF*~
```

FIGURE 8-3

(a)



(b)

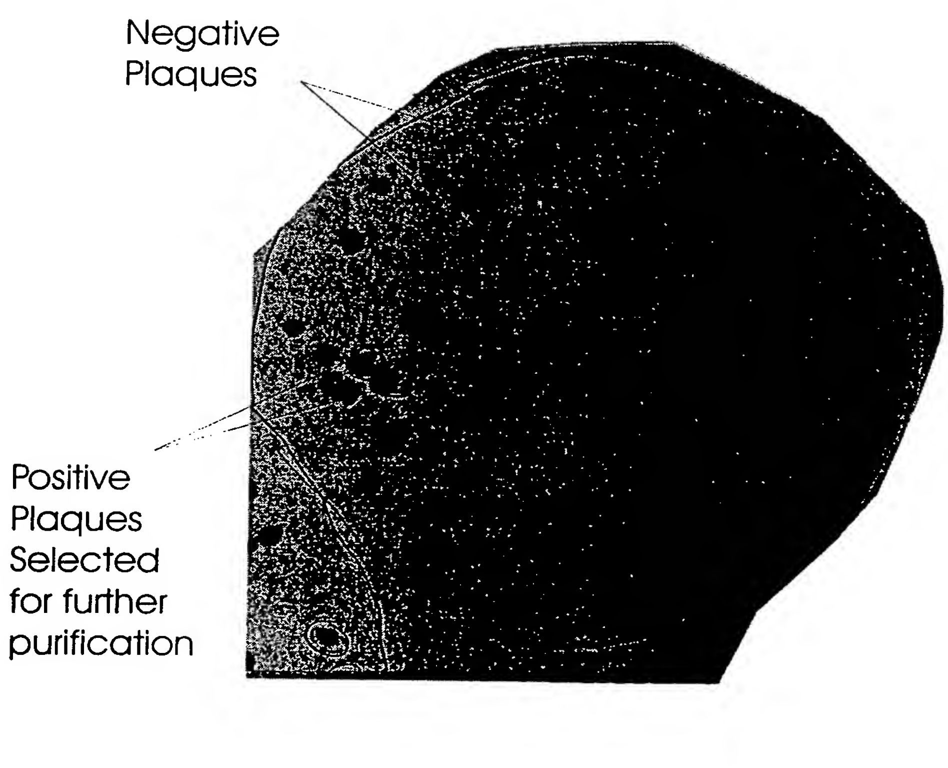
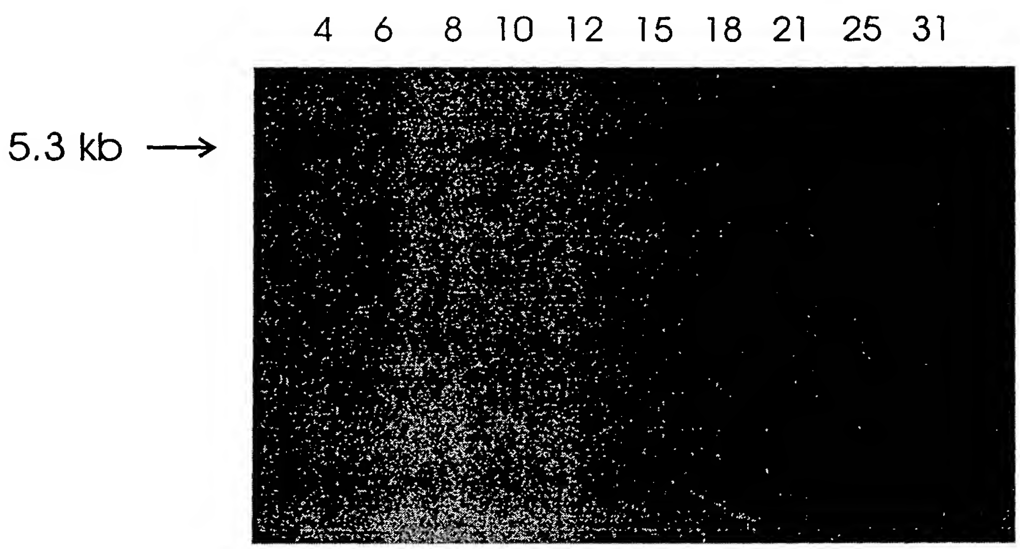
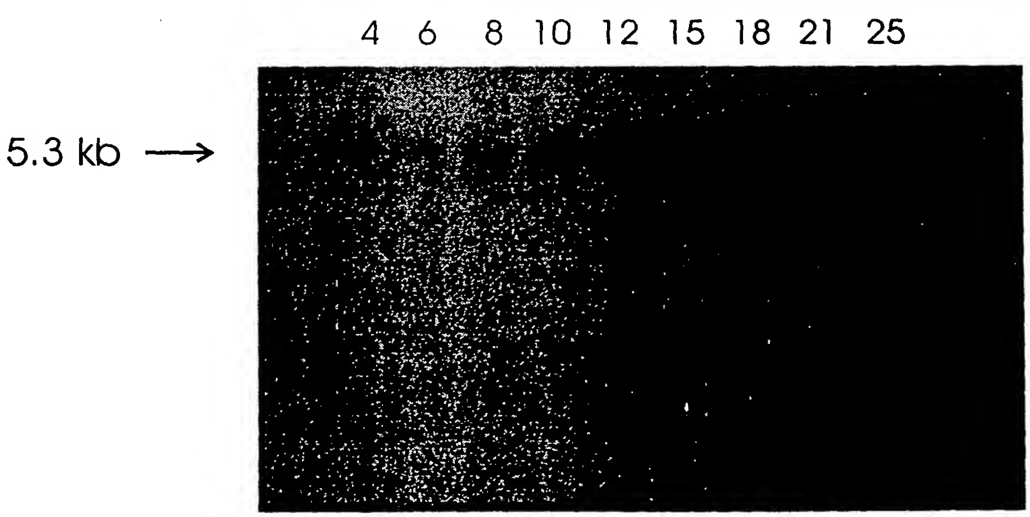

Negative Plaques
Positive Plaques Selected for further purification

FIGURE 9

(a) Wyuna



(a) Gabo



(C) Gabo

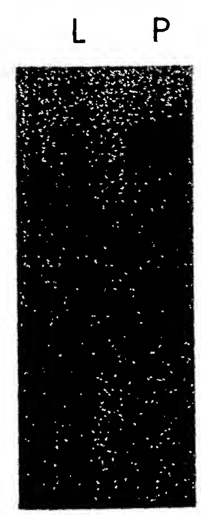


FIGURE 10

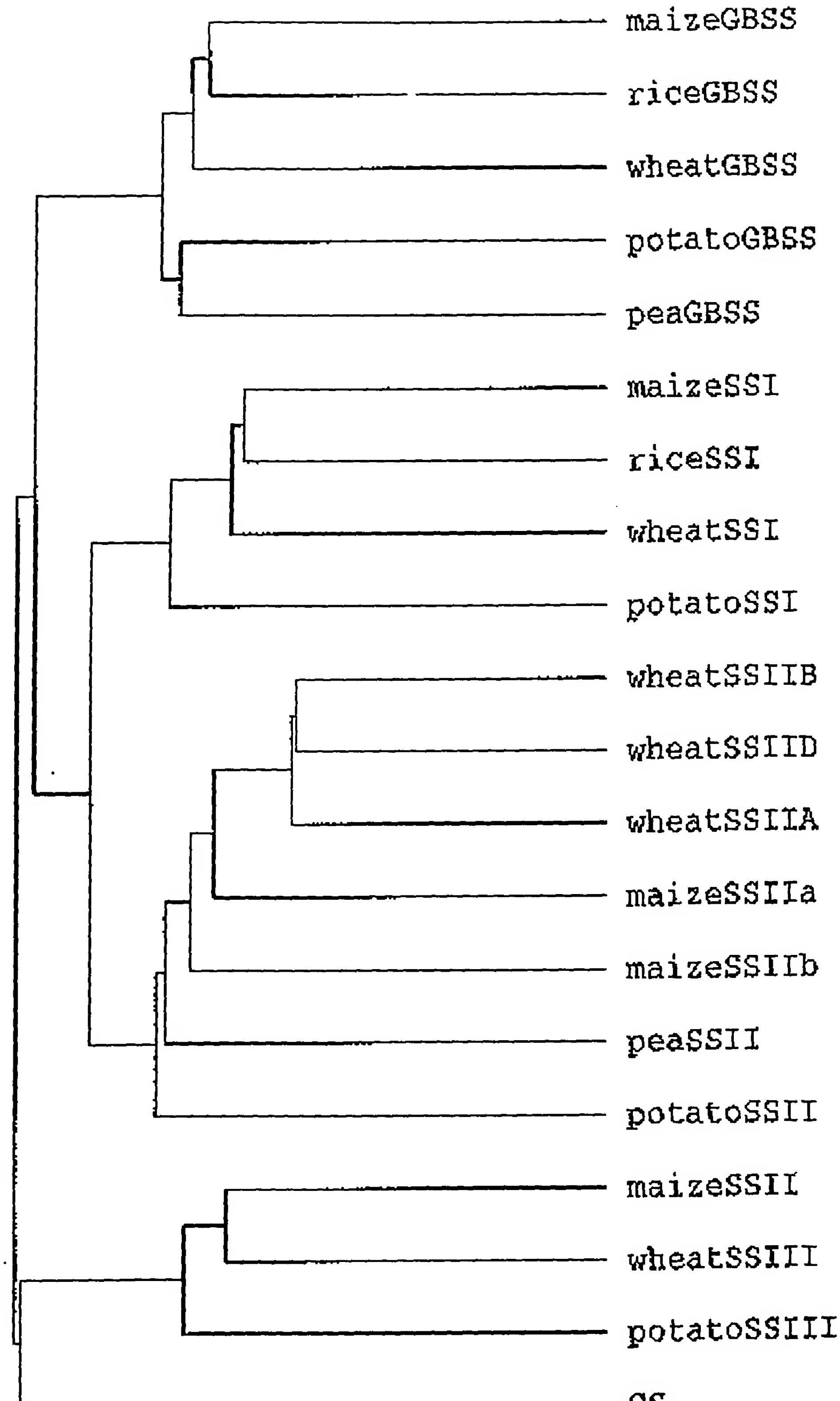

maizeGBSS
riceGBSS
wheatGBSS
potatoGBSS
peaGBSS
maizeSSI
riceSSI
wheatSSI
potatoSSI
wheatSSIIB
wheatSSIID
wheatSSIIA
maizeSSIIa
maizeSSIIb
peaSSII
potatoSSII
maizeSSII
wheatSSIII
potatoSSIII
GS

FIGURE 11

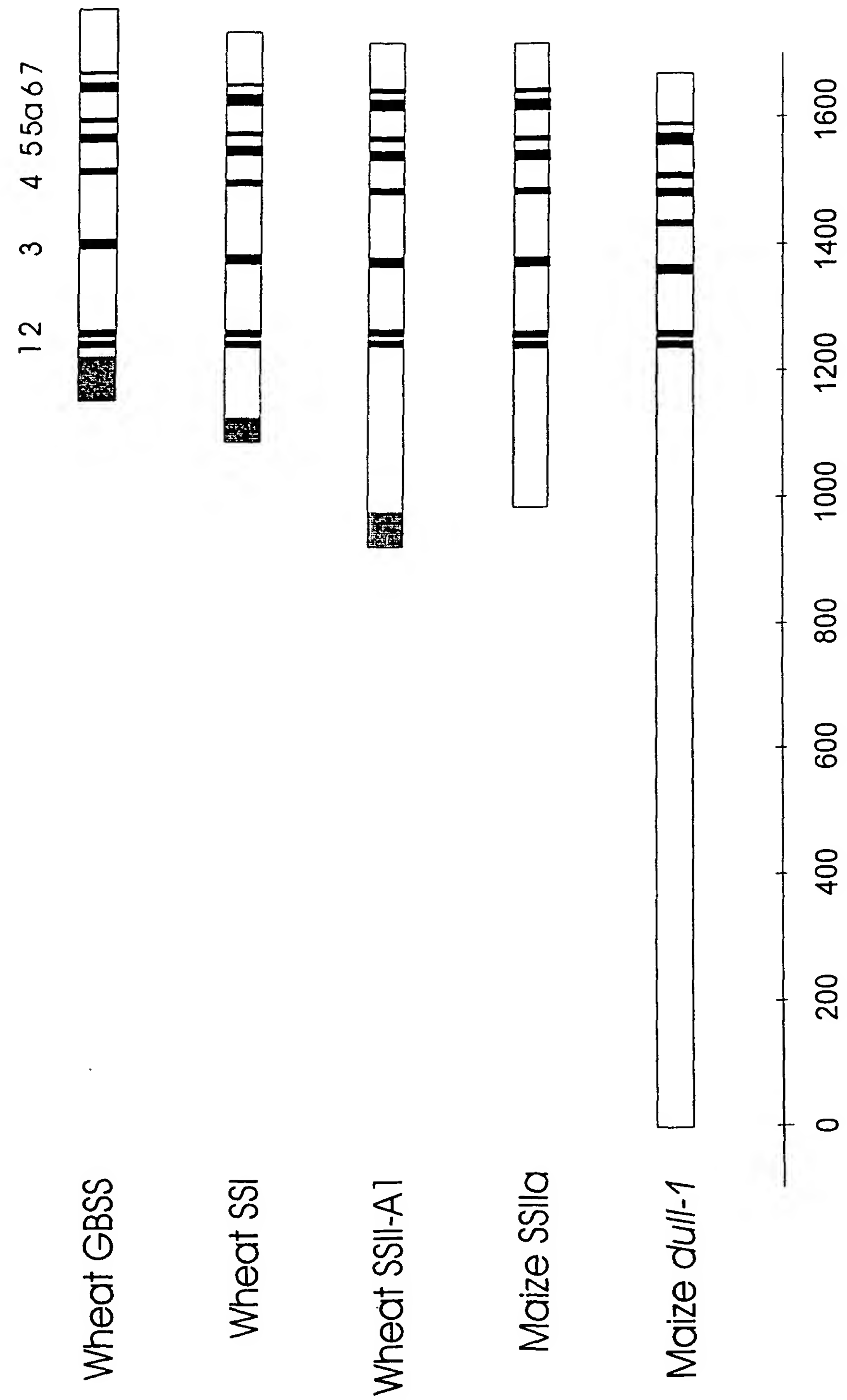

1 2 3 4 5 5a 6 7
Wheat GBSS
Wheat SSI
Wheat SSII-A1
Maize SSIIa
Maize dull-1
0
200
400
600
800
1000
1200
1400
1600

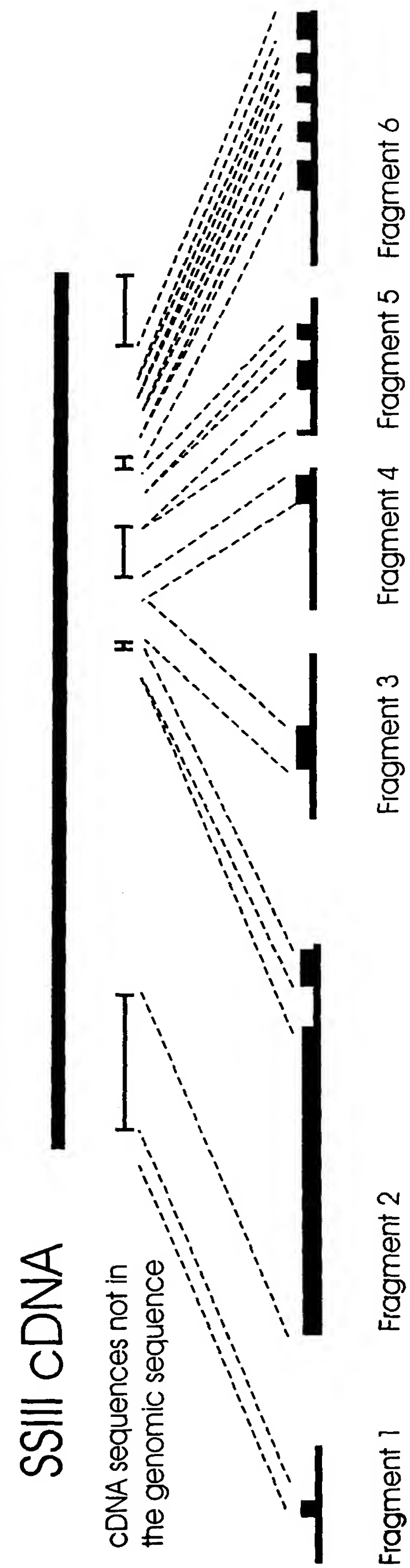
FIGURE 12

```
                      Region 1            Region 2
             10         20         30         40         50         60         70         80         90
wGBSS    81 FVGAEMAPWS KTGGLGDLLG GLPPAMAANG HRVMVISPRY DQYKDAWDT- -----SVVSE IKVVDKYERV RYFHCYKRGV DRVFVDHPCF   170
wSS1    144 -*TG*A**YA *S*****VC* S**I*L**R* *****VM*** LNGSSDKNYA KALYTGKHIK *PCFGGSHE* TF**E*RDN* *W******SY   233
wSS2    314 --A**CS**C *******VA* A**K*L*KR* *****VV*** GD*EE*Y*V- -----G*RKY Y*AAGQDME* N***A*ID** *F**I*A*L*   403
wSS3   1187 -IAV****VA *V*****VVT S*SR*IQDL* *T*E**L*K* *CLNQSSVK- ---------- -DLHLYQSFS WGGTEI*VW* G**EDLTVY*  1276

                                                                             Region 3
            100        110        120        130        140        150        160        170        180
wGBSS   171 LEKVRGKTKE KIYGPDAGTD YEDNQQRFSL LCQAALEVPR ILNLDNNPYF SGPYGEDVVF VCNDWHTGLL ACYLKSNYQS NGIYRAAKVA   260
wSS1    234 -HRPGSLYGD ------NFGA FG***F*YT* **Y**C*A*L **E*GGYI*G QN-----CM* *V****AS*V PVL*AAK*RP Y*V**DSRST   323
wSS2    404 RHRQEDIYGG ---------S RQEIMK*MI* F*K**V***W HVPCGGV**G D*----NL** IA*****A** PV***AY*RD H*LMQYTRSI   493
wSS3   1277 **PQN*MFGV ------GCVY GRNDDR**GF F*HS***--F **QNEFS*H- --------II H*H**SSAPV *WLY*EH*SQ -SRMASTR*V  1366

            190        200        210        220        230        240        250        260        270
wGBSS   261 FCIHNISYQG RFSFDDFAQL NLPD-----R FKSSFDFIDG YDKPVEGRKI NWMKAGILQA DKVLTVSPYY AEELISGEAR GCELDNIMRL   350
wSS1    324 LV***LAH** LEPASTYPD* G**PEWYGAL EWVFPEWARR HALDKGEAVN FLKG*VVTAD RI*TVSQG*S W*VTTAEGGQ *LNELLSS*K   413
wSS2    494 MV****AH** *GPV*E*PFT E**------- -EHYLEHFRL **PVGGEHAN YFAAGLKMAD QV*VVSPG*L W*LKTVEGGW *LHDIIRQND   583
wSS3   1367 *T***L-EF* AHYIGKAMTY CDK------- ---------- ---------- ---------- --------AT TVSPTYSRDV AGHGAIAPHR  1456

              Region 4                                                           Region 5
            280        290        300        310        320        330        340        350        360
wGBSS   351 TGITTIVNGM DVSEWDPTKD KFLAVNYDIT TALEGKALNK EALEGKALNK EALQAEVGLP VDRKVPLVAF IGRLEEQKGP DVMIASIPEI   440
wSS1    414 SVLNG****I *IND*N**T* *C*PHH*SV- ---------- DD*S***KC* AE**K*L*** *RED***IG* ****DY***I *LIKMA***-   503
wSS2    584 WKTRG****I *NM**N*EV* VH*KSDGYTN ------FSLG TLDS**RQC* ****R*L**Q *PAD***LG* ****DG***V EIIADAM*W*   673
wSS3   1457 EKFYG*L**I *PDI***YT* N*IP*P*TCE ----NVVEG* **AKRALQQ* FG**QT---- ---D**I*GI *T**TA***I -HL*KHAIHR  1546

             Region 5a                                                 Region 6                          Region 7
            370        380        390        400        410        420        430        440        450
wGBSS   441 LKEEDVQIVL LGTGKKKFER LLKSIEEKFP SKVRAVVRFN ------APLA HQMMAGADVL AVTSRFEPCG LIQLQGMRYG TPCACASTGG   530
wSS1    504 *MR****F*M **S*DPI**G WMR*T*SSYK D*F*GW*G*S ------V*VS *RIT**C*I* LMP******* *N**YA*Q** *VPVVHG***   593
wSS2    674 V-SQ***L*M ****RHDL*S M*RHF*REHH D***GW*G*S ------VR** *RIT****A* LMP******* *N**YA*A** *VPVVHAV**   763
wSS3   1547 TL*SNG*V** **SAPDHRIQ GDFCRLADAL HG*YHGRVKL -VLTYDE**S *LIY**S*FI I*P*I***** *T**VA**** SIPIVRK***  1636
```

FIGURE 13-1

```
        Region 7 (Continued)
          460        470        480        490        500        510        520        530        540
wGBSS 531 LVDTIVEGKT GFHMGRLSYD CNVVEPADVK KVVTTLKRAV KVVGTPAYHE MVKNCMIQDL SWKGPAKNWE DVLLELGVEG SEPGIVGEEI  620
wSS1  594 *R**-**TFN ---------- --PFGAKGEE GTGWAFSPLT VDKMLW*LRT AMSTFREHKP **E*LM*RGM TKDHTWDHAA EQYEQIF*WA  683
wSS2  764 VR**-*PPFD ---------- --PFNHSGLG ---W*FD**E AHKLIE*LGH CLRTYRDYKE **R*LQERGM SQDFSWEHAA KLYED*LLKA  853
wSS3 1637 ****-*FDVD NDKDRAR*LG LEPNGFSFDG ADSNGVDY*L NRAIGAWFDA RDWFHSLCKR VMEQDWSWNR PA*DYIELYH AARKF*....  1726

550       560        570        580        590        600        610        620        630
wGBSS 621 APLAMENVAA P*........ .......... .......... .......... .......... .......... .......... ..........  710
wSS1  684 FVDQPYVM.. .......... .......... .......... .......... .......... .......... .......... ..........  773
wSS2  854 KYQW...... .......... .......... .......... .......... .......... .......... .......... ..........  943
wSS3 1727 .......... .......... .......... .......... .......... .......... .......... .......... ..........  1816
```

FIGURE 13-2



SSIII cDNA
cDNA sequences not in the genomic sequence
Fragment 1
Fragment 2
Fragment 3
Fragment 4
Fragment 5
Fragment 6

FIGURE 14